République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique
Université A.MIRA-BEJAIA

Faculté des Sciences Exactes
Département de Mathématiques
Laboratoire de Mathématiques Appliqués

THÈSE

EN VUE DE L'OBTENTION DU DIPLOME DE DOCTORAT

Domaine : Mathématiques et informatique Filière : Mathématiques
Spécialité : Probabilités et Statistiques

Présentée par
BENAKLEF Nesrine

Thème

Etude des processus fractionnaires à coefficients périodiques

Soutenue le: 29/10/2025 **Devant le Jury composé de:**

Nom et Prénom	Grade		
Mme ZEROUATI Halima	**Professeur**	U. Bejaia	Présidente
Mme BELAIDE Karima	**Professeur**	U. Bejaia	Rapporteur
Mme SAGGOU Hafida	**Professeur**	USTHB	Examinatrice
Mr HAMAZ Abdelghani	**Professeur**	U. Tizi Ouzou	Examinateur
Mme BARECHE Aicha	**Professeur**	U. Bejaia	Examinatrice
Mme OURBIH Megdouda	**Professeur**	C.U. Tipaza	Invitée

Année Universitaire: 2025-2026

People's Democratic Republic of Algeria
Ministry of Higher Education and Scientific Research
A.MIRA-BEJAIA University

Faculty of Exact Sciences
Department of Mathematics
Laboratory of Applied Mathematics

THESIS

IN VIEW OF OBTAINING A DOCTORATE DEGREE

Domain: Mathematics and Computer Science   Field: Mathematics
Specialty: Probability and Statistics

Presented by
BENAKLEF Nesrine

Theme

Study of fractional processes with periodic coefficients

Defended on: 29/10/2025 **Before the Jury composed of:**

First and last name	Title		
Mrs ZEROUATI Halima	Professor	U. of Bejaia	Chairwoman
Mrs BELAIDE Karima	Professor	U. of Bejaia	Supervisor
Mrs SAGGOU Hafida	Professor	USTHB	Examiner
Mr HAMAZ Abdelghani	Professor	U. of Tizi Ouzou	Examiner
Mrs BARECHE Aicha	Professor	U. of Bejaia	Examiner
Mrs OURBIH Megdouda	Professor	U.C. of Tipaza	Invited

Academic Year: 2025-2026

Acknowledgements

In the name of Allah, the Compassionate, the Merciful

{I only intend reform to the best of my ability. My success lies with Allah. In Him I trust; and to Him, I turn}

God Almighty has spoken the truth

[Hud from verse:88]

First and foremost, I would like to express my heartfelt gratitude to my supervisor, Professor K. Belaide, for her unwavering support and encouragement throughout the entire process of working on my thesis. Without her guidance, completing this work would not have been possible. Her compassion, patience, helpfulness, and availability kept my passion for learning and research alive. Words cannot fully convey how thankful I am to her. Thank you for being a steadfast source of guidance and inspiration.

I wish to extend my heartfelt thanks to Professor H. Zerouati from the University of Bejaia for graciously accepting to chair the jury of this thesis.

I am profoundly grateful to Professor H. Saggou from USTHB, Professors A. Bareche and M. Ourbih from the University of Bejaia, as well as Professor A. Hamaz from the University of Tizi Ouzou, for their willingness to serve on the jury of my thesis.

I am forever grateful to my parents for their boundless love and support, which have been the pillars of my strength throughout this journey. To my late father, whose memory continues to inspire me, I am deeply thankful.

I would like to express my profound gratitude to A. D., a dear friend of mine.

I am also deeply thankful to my family, whose unconditional love, patience, and trust have always been my greatest support. My heartfelt thanks go as well to my friends and colleagues for their kindness, understanding, and constant encouragement along the way.

This thesis is dedicated to

The memory of my father, my first mentor and my final guide.

My mother, the greatest woman in my world.

Myself,
for the persistence and determination it took to reach this milestone.

To all who prayed for me.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF WORKS

Publications

• Fractional periodic autoregression. (Published in Stochastics: An International Journal of Probability and Stochastic Processes).

• Minimum Hellinger distance estimates for a periodic fractional autoregressive model (Published in Communication in Statistics - Simulation and Computation)

• On several local asymptotic properties for periodic fractional autoregressive models (Submitted).

Communications

• On fractional autoregressive model of order 1 with a periodic coefficient. 1st International Conference on Pure and Applied Mathematics IC-PAM'S21, May 26-27, 2021, Ouargla, Algeria.

• Local asymptotic normality for FAR(1) with a periodic coefficeint. Fourth Edition of the International Conference on Reasearch in Applied Mathematics and Computer Science ICRAMCS 2022, March 24-26, 2022, Morocco.

• On local asymptotic properties for fractional autoregressive model with a periodic coefficient. Conférence nationale Statistique, Modélisation et Estimation Paramétrique et Non Paramétrique SMEPNP'22, Novembre 06-09, 2022, Bejaia, Algerie.

• LAN property related efficient notions for FAR(1) model. Deuxième Conférence nationale des Mathématiques Pures et Appliquées CMPA 2023, 04 Juillet 2023, Tébessa, Algerie.

• On fractional processes. The twelfth Meeting on Mathematical Analysis and Applications RAMA'12, November 09-12, 2023, Adrar, Algeria.

• Estimation pour les modèles autoregressifs fractionnaires d'ordre 1 à coefficients constants, December 13-14, 2023, Guelma, Algeria.

• Hybrid model for modelling crude oil production in Algeria. Statistiques et Analyse Avancées: Domaines d'Interactions et d'Applications 1 SAADIA 1, 20-22 Novembre 2023, Bejaia, Algerie.

• Estimation in short-long memory time series. Statistiques et Analyse Avancées: Domaines d'Interactions et d'Applications 1 SAADIA 1, 20-22 Novembre 2023, Bejaia, Algerie.

• First-order fractional autoregressive models. The First National Conference on Differential Geometry and Dynamical Systems DGDS 2023, December 19-20, 2023, Relizane, Algeria.

• Estimation for periodic FAR(1) model: theory and simulation. First National Conference on Non Linear Differential Equations and thier Application NCNDEA'2024, February 10-11, 2024, Oum El Bouaghi, Algeria.

• MHD estimates for classic FAR(1) models, theory and simulation. The 1st International Conference on Nonlinear Mathematical Analysis and its Applications IC-NMAA'24, May 14-15, 2024, Bordj Bou Arréridj, Algeia.

NOTATIONS AND SYMBOLES

Symbol	Definition
$\mathbb{E}$	Mathematical expectation.
$\mathbb{P}$	Probability.
$\Gamma(\cdot)$	Gamma function.
$F(.,.,.)$	Hypergeometric function.
I_f	Fisher information.
$\mathbb{Z}$	Set of integers .
$\mathbb{N}$	Set of natural numbers.
$[x]$	Floor function.
$\equiv$	Congruence relation.
$\sim$	Asymptotic equivalence.
LAN	Local asymptotic normality .
LAQ	Locally asymptotically quadratic.
LAM	Local asymptotic minimaxity.
H^2	Hellinger distance.
$MHDE$	Minimum Hellinger distance estimate.
CSS	Conditional sum of squares.

GENERAL INTRODUCTION

0.1 The evolution of time series analysis: A historical perspective

Understanding how phenomena vary over time is a fundamental issue in many scientific disciplines and requires tools to capture and model temporal dependencies. One of the most common frameworks for studying temporal evolution is time series analysis. A time series is a sequence of observations taken sequentially in time that allows researchers to model, understand and forecast changes in various phenomena. In other words, the theory of time series is primarily concerned with constructing a plausible mathematical model for a data set, studying its main characteristics, and using it to make accurate predictions in the most effective way. Many datasets are in the form of time series: a monthly record of sales, a weekly count of hospital admissions, daily temperature readings and hourly measurements of electricity consumption, and so on. Time series are prevalent in fields such as economics, business, engineering, natural sciences (particularly geophysics and meteorology), and social sciences. An intrinsic feature of a time series is that successive observations often show patterns or trends over time.

Time series analysis focuses on techniques for identifying and analyzing these trends. It involves the development and application of stochastic processes (models that describe the probability structure of such sequences).

The study of time series analysis has a long history, beginning centuries ago with the need to understand patterns in sequential data collected over time. One of the earliest contributions came from Graunt in the 1600s, who analysed death records that had been archived in London since the early 1500s [79]. Though his work was not mathematically advanced, it laid the groundwork for what would become a systematic study of time series data. Building on this foundation, Fourier's contributions [38] in the 19th century, specifically his work on harmonic analysis, provided essential mathematical tools for studying periodic patterns. Fourier's approach, which involved decomposing complex cycles into simpler components, has since become a fundamental aspect of time series analysis.

By the late 19th century, researchers such as Thiele were developing formal statistical methods for analyzing time series. Thiele's work [68],[69] introduced key concepts such as noise and randomness, which are still central to modern approaches. At around the same time, studies on phenomena such as Chandler's wobble [56] (a periodic motion of the Earth's poles) highlighted the importance of time series in understanding natural phenomena.

At early in the 20th century, mathematicians like Yule [112] were among the first to apply autoregressive (AR) models to time series data, a significant step forward in the field of formal statistical modelling. These models enabled the analysis and forecasting of trends, influencing numerous fields, including economics and engineering. Shortly thereafter, Walker [105] modified Yule's method, leading to the development of the Yule-Walker equations, which are used to determine the order of the autoregressive process in a time series.

The development of time series analysis accelerated in the mid-20th century. A significant milestone was the publication of "Time Series Analysis: Forecasting and Control" by Box and Jenkins [10] in 1970. This book introduced a systematic approach to identifying, estimating, and validating models, popularizing the autoregressive integrated moving average (ARIMA) framework for analyzing both stationary and non-stationary series. Their work also formalized seasonal ARIMA (SARIMA) models, which incorporate ARIMA's core principles with explicit seasonal components. Simultaneously, advances such as the Kalman [60] filter have provided effective methods for dealing with noisy and missing data (Jones [57]), thereby significantly improving the capacity to extract valuable information from complex datasets.

The second half of the 20th century also saw the rise of non-linear models and methods to deal with long-range dependence in time series data. Long-range dependence, or long-memory, characterizes series in which correlations between values decay slowly over time, resulting in persistent effects. Granger and Joyeux (1980) [37] and Hosking (1981) [50] introduced fractional differencing as a way to model such phenomena, culminating in the development of the autoregressive fractionally integrated moving average (ARFIMA) model. These models expanded the analytical framework, enabling researchers to better study and predict time series with long-term dependencies. In addition, non-linear models addressed limitations of linear approaches by capturing more complex forms such as sudden shifts and asymmetries in data. Similarly, Granger's work [35] on causality provided tools to explore relationships between variables. The creation of methods such as fractal analysis [46], [47] and wavelet transforms [91] has led to the availability of a wider range of tools, facilitating the analysis of irregular data structures.

As computing power advanced in the late 20th century, it had a considerable impact on time series analysis, with the application of the Kalman filter [45],[63] becoming more frequent. In addition, Markov Chain Monte Carlo (MCMC) methods, especially Gibbs sampling [29], have provided powerful tools for analyzing non-linear and non-Gaussian models, including applications in time series analysis. These approaches were revolutionary as they offered robust frameworks for addressing such complexities and significantly advanced simulation techniques. The MCMC method facilitates the generation of simulations that provide deeper analysis of time series data.

Another noteworthy development during this period was the integration of multivariate time series analysis, with the need to consider jointly several related time series [85]. One critical development was the emergence of vector autoregressive (VAR) models, while another pivotal contribution originally introduced by Engle and Granger [23] was cointegration analysis, which examined long-term relationships between variables.

In the 21st century, advancements in data collection and computational tools have resulted in further transformation of time series analysis. With the increasing availability of massive datasets, researchers have focused on developing more sophisticated mathematical tools. Improved techniques for detecting and analyzing complexity in time series have arisen. Recurrence plots and multi-fractality analysis are tools that allow analysts to understand the underlying structure of complicated systems, which is crucial for forecasting in areas such as weather prediction and financial markets. Additionally, methods such as delay vector variance (DVV) [28],[2] provide insights into the ability to predict non-linear systems.

Recent years have also seen the introduction of non-mathematical frameworks for the application of time series analysis. Machine learning [55],[14] and artificial intelligence [106] have been incorporated into time series analysis. These approaches are particularly useful for modelling highly complicated time series systems. Neural networks and other advanced algorithms are being used more and more to predict and detect patterns in big data sets , complementing traditional mathematical methods and expanding the range of time series applications.

An additional novel technique that has seen considerable popularity in recent years is hybrid approaches [93], which combine traditional time series methods with artificial intelligence techniques. (Further details on the evolution of time series analysis are well described in [102] and [30]).

0.2 Motivation and background of the study

In 1951, Hurst [53] studied fluctuations in the water levels of the Nile to understand the long-term behavior of hydrological time series. Through his work, he found that the fluctuations do not follow a completely random structure, but rather show long-term persistence.

In 1981, Hosking [50] extended Hurst's work by developing and applying the ARFIMA model. In 1984 [51], he studied the time series of annual flows in the Nile and used the fractional differentiation method to model the long-term persistence present in this hydrological series. Hosking applied several fractional processes, including processes satisfying the following stochastic equation

$$(1 - L)^d X_t = \varepsilon_t, \tag{1}$$

with L as a lag operator. $\{\varepsilon_t, t \in \mathbb{Z}\}$ is a strong white noise, a sequence of independent random variables identically distributed with zero mean and finite variance.

By plotting the temperature data series for Bejaia (Figure 1, the data were collected as daily average temperatures measured 2 meters above the ground surface and expressed in degrees Celsius. They are available on the following website: Bejaia temperature), covering the period from January 1^{st} 2020 to March 31^{st} 2025, we found that the autocovariance function shows a slow decay, suggesting that a long-memory model, such as the one used by Hosking in 1984, may be suitable for modelling the data series. However, the graph of the data also shows a recurrent form or a form of periodicity that is not captured by the classical ARFIMA model. In light of this, we propose to include the periodicity in the stochastic equation presented in (1).

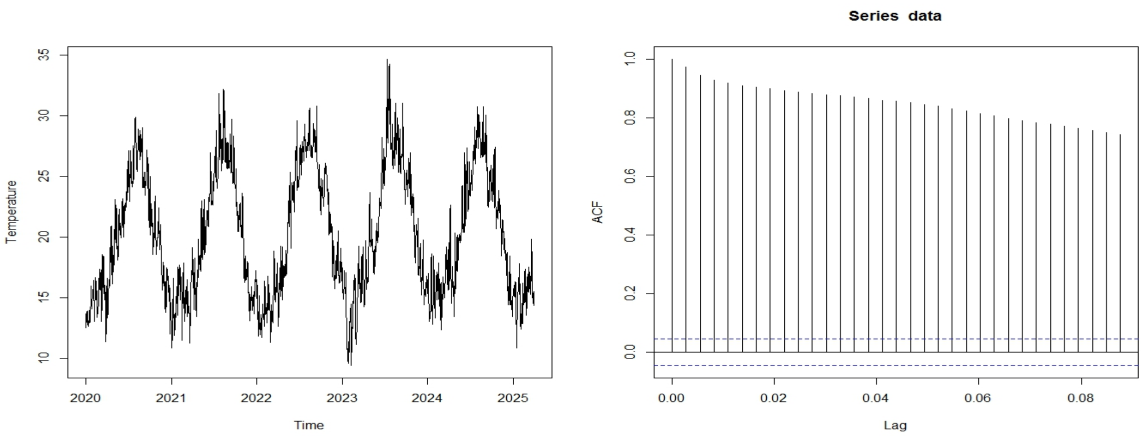

Figure 1: Representation of the temperature data set of Bejaia.

Stochastic processes were developed in the early 20th century through the work of

pioneers such as Kolmogorov [64], who formalized the mathematical framework of probability and stochastic processes, and Ross [89], whose contributions in the 1970s and early 1980s expanded the understanding and application of these processes.
The first application of stochastic processes goes back to the modelling of Brownian motion, a phenomenon observed by Robert Brown in 1827 and mathematically formalized by Einstein [22] and Smoluchowski [96] in 1905 and 1906.
Building on foundational work in the early 20th century [113], stochastic processes have since found extensive application in time series analysis. An important class of these models are stationary models, which assume that the process remains in statistical equilibrium (a state in which statistical properties such as mean and variance do not change over time). These models are characterized by a constant mean and variance.
Early contributions to stationary processes include Yule's pioneering development of autoregressive modeling in 1927, Walker's (1931) [105] generalizations, and Slutsky's (1937) [95] demonstration of how deterministic transformations can produce apparently random series. The Wold decomposition theorem, introduced in 1938 [108], provided a theoretical foundation by showing that any stationary process can be expressed as the sum of deterministic and stochastic components. Subsequent advancements by Kendall (1945) [62] and Bartlett (1946) [7] in autocorrelation structures and spectral analysis deepened the understanding of stationary models. Quenouille (1952 [84], 1957 [85]) refined parameter estimation techniques, while Doob (1953) [12] established the stochastic basis of stationarity. Grenander and Rosenblatt (1957) [39] expanded spectral theory, Hannan (1960) [44] developed methods for model identification.

Despite their usefulness, stationary models often fail to reflect the behavior of real-world time series, particularly in practical applications where non-stationary processes, characterized by the absence of a stable mean level, are common. In this context, forecasting methods developed by Holt (1957) [49], Winters (1960) [107], and Brown (1962) [13], which utilise exponentially weighted moving averages (EWMA), have proven particularly effective for addressing certain types of non-stationary processes. The EWMA forecast, as demonstrated to minimize the mean square error Muth (1960)[77], is related to a wider class of non-stationary processes known as autoregressive integrated moving average (ARIMA) processes. These processes include stationary and non-stationary models, offering flexible models for the diverse behaviors captured in practice.

ARIMA models, which gained widespread recognition with Box and Jenkins in 1971 [10], were explored earlier by Yaglom in 1955 [111]. These models are particularly useful for analyzing non-stationary time series, as they incorporate both autoregressive (AR) and moving average (MA) components, along with differencing to handle non-stationarity. The concept of differencing itself is not new, having been part of earlier methods like the variate difference technique outlined by Tintner (1940) [100] and Rao and Tintner (1963) [87], which laid the groundwork for later developments. Additionally, non-stationary processes were studied by Zadeh and Ragazzini (1950) [114], and Kalman's work in the 1960s [59], [60] on filtering and control processes made important contributions to the modelling of such time series.

Most models developed prior to 1980, despite their differences, shared a common limitation: their predictions were only effective over short time horizons. Consequently, these models were characterized by short-range dependence, where correlations between observations diminished rapidly as the time lag increased. And here was the appearance of the models that motivated this study, developed to handle what is known as long-range dependence in a serie.

Long-range dependence is a widely applicable concept that extends beyond hydrological phenomena, with notable applications including seismic data, as developed by Ogata and Abe (1991) [81], and meteorological data, studied by Reisen and Lopes (1999) [88]. In telecommunications and electricity, Soares and Souza (2006) [97] demonstrated its relevance, while Ferrara and Guegan (2000) [25] identified long-memory components in monthly passenger traffic data from the Paris transport network (RATP). In the energy sector, Diongue (2005) [20] used long-memory models to analyze electricity prices under the liberalization of the European market, observing price spikes, mean reversion, and non-constant volatility. Atkins and Chen (2002) [5] initially attempted to model Canadian electricity prices with short-memory processes, but unit root tests revealed non-stationarity. To address this, they employed ARFIMA processes, which proved effective.

Economics and finance remain central areas for long-memory process applications. Studies have found long-memory components in inflation rates (Baillie, 1996 [6]), stock prices (Cheung and Lai, 1995 [16]), exchange rates and GDP series (Cheung, 1993 [15]).

0.3 An overview of the thesis

The present thesis thoroughly addresses a time series model that extends the fractional autoregressive model (FAR) with a constant coefficient to the periodic fractional autoregressive model.

The first chapter introduces fundamental concepts relevant to the study. The second chapter of the thesis presents the model and its initial parameters. We detail the forms of causality and invertibility that allow us to derive their conditions based on the multivariate process representation related to the original periodic univariate process alongside the results of Serroukh (1996) [94]. This approach allows us to analyze the stationary form of the process, as the variance function reveals non-stationarity, leading us to demonstrate that it is periodically correlated. Thus, we establish a multivariate stationary process that corresponds to the periodic univariate non-stationary process. It is important to note that Gonçalves (1988) [34] conducted a comprehensive study of FAR models with constant coefficients, as well as long-memory processes, in her thesis. Later, Serroukh (1996) [94] expanded upon this work by addressing additional aspects of time series theory for FAR models, thereby contributing to a more complete framework. However, both studies were limited to models with constant coefficients, whereas our work extends beyond this restriction.

Periodically correlated processes, also known as cyclostationary processes, are a class of stochastic processes characterized by statistical properties, such as the mean and autocovariance function, that vary periodically over time. The study of these processes originates from the pioneering work of Gladyshev (1961) [31], who formalized their mathematical framework, with further contributions made in (1963) [32]. Hurd (1970) [52] advanced this foundation by providing important theoretical insights into their structure and properties. Later, Gardner (1994) [104] significantly expanded the field by demonstrating their practical applications, particularly in signal processing and communications. Building on this series of papers, Boshnakov (1994) [9] developed spectral representations of periodically correlated processes, while Trootman (1979) [101] developed methods for estimating periodic autocovariance functions. Cipra (1985) [19] further applied these processes to time series forecasting, particularly for seasonal or cyclical patterns.

We present the general forms of the autocovariance and autocorrelation functions, as well as their asymptotic representations. These functions require further detail, as their final forms are expressed in terms of the periodic coefficients we introduce. To address this, we provide additional results in this context. Furthermore, we prove that our process satisfies one of Gauss's contiguous relations. These relations were explored by Gonçalves (1988) [34], where she constructed an estimator based on them.

In the third chapter, we spotlight estimation, which is a crucial notion in time series. Although the presence of periodicity and non-stationarity widens the applicability of time series in modeling several phenomena, on the other hand, it affects different aspects of the theory, particularly the theory of estimation; it restricts the applicability of traditional methodologies such as moments estimation and maximum likelihood estimation. These techniques were considered in parameter estimation of the FAR with constant coefficient by Gonçalves (1988) [34].

The non-stationarity property even affects alternative techniques such as Whittle likelihood estimation due to the lack of a closed-form expression of spectral density. To remedy this, we propose to construct the estimate for the multivariate stationary process related to the initial non-stationary. We propose estimating by minimizing the Hellinger distance, which was used for the first time in estimation for parametric models by Beran in 1977 [8]. Since then, this theory has been a framework around which many works have been published. We cite among them Ndongo et al. (2015) [78], where they compared the conditional sum of squares and minimum Hellinger distance procedures considering ARFIMA with infinite variance innovations; Mbeke and Hili (2018) [75] developed an MHDE for multivariate stationary ARFIMA. More recently, Amimour and Belaide (2022) [4] used the technique on a periodically time-varying long-memory parameter. In our study, in addition to constructing the MHDE, we also investigate its consistency and asymptotic distribution under some regularity assumption. To enrich the study we propose another estimation technique proven to be equivalente to the maximum likelihood estimates by Box and Jenkins (1970) [10] known as conditional sum of squares (CSS) estimates, and we conduct a study to compare both estmations.

In the fourth chapter we provide local asymptotic properties, which are the local asymptotic normality (LAN), local asymptotic linearity and the local asymptotic minimaxity (LAM). We use the theory developed by Swensen (1985) [98] but initially introduced by Le Cam in 1960, who devoted a series of research studies to it, among which we mention [70], where he created the LAN general theorem and [71], where additional theories related to asymptotic problems in statistics were thoroughly studied.

The fundamental concept of LAN is that the log-likelihood ratio is locally asymptotically normally distributed, with a locally linear mean of the parameters of the statistical model under study and a locally constant variance. Several criteria were established to prove that a statistical model has the LAN property. Raoussas (1979) [90] searched for the conditions under which the LAN property can be obtained for stochastic processes; later, Swensen (1985) [98] developed that theory for autoregressive time series with a regression trend. A worth noting point is that Swensen extended the Le Cam lemma to an easier-to-apply lemma, and based on the extended lemma, we prove the LAN property in this study.

Among several authors who discussed LAN in periodic models, we make reference to Garel and Hallin (1995) [26], who proved the LAN property for multivariate general linear models with an ARMA error term. Choy et al. (1999) [17] proposed the asymptotic theory for a regression model with stationary non-Gaussian ARFIMA(p,d,q) long-memory errors based on the LAN method. Moreover, Amimour and Belaide (2022) [3] established the LAN for a periodically-time varying

long-memory parameter model.
This concept has been a popular research subject that provided a prerequisite for constructing locally asymptotically minimax (LAM) estimators that were firstly described by Hàjek (1972) [42] and subsequently studied starting by Fabian and Hannan (1982) [24], who established the LAM for a family of models that satisfy the LAN. On the other side, Kreiss (1987) [67] established and researched the first effective factors to take into account when estimating time series models and he introduced the LAN property for the ARMA process, without ignoring the pioneering work of Serroukh (1996) [94], where all the asymptotic properties we are interested in were extensively studied, Haddad and Belaide (2020) [41] investigated the LAN property for long-memory processes with strong mixing noises. Later, Seba and Belaide (2024) [92] extended their work and established other local asymptotic theories related to the LAN property.

The last three chapters were completed with simulation studies applied to the key results of each chapter using R software. Notably, in the simulation section for the MHDE, we constructed a confidence interval and presented it graphically to demonstrate the estimator's efficiency.
We conclude by summarizing our main findings and suggesting valuable avenues for further studies.

CHAPTER 1

GENERALITIES

After the introduction, this chapter summarizes some key definitions and properties necessary to understand the topics studied in the subsequent sections. We begins with the introduction of fundamental definitions and mathematical tools. Next, we present the concept of long-memory processes. Finally, we touch on the Hellinger estimator and the concept of LAN.

1.1 Stochastic processes and periodicity

Definition 1.1 (Stochastic process). [54] A stochastic process is a collection of random variables $\{X_t, t \in T\}$ defined on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$, where T is an index set (usually representing time). For each fixed $t \in T$, X_t is a random variable, and for each fixed $\omega \in \Omega$, $\{X_t(\omega), t \in T\}$ is a realization or path of the process.

Definition 1.2 (Time series process). [83] A time series process is a stochastic process $\{X_t, t \in T\}$, where T is a set of time indices (typically $T \subseteq \mathbb{Z}$), and X_t is a random variable representing the observation at time t. The process is characterized by its temporal ordering, and the dependence structure between observations at different times is often of primary interest. In practice, a time series is observed as a finite sequence $\{X_1, X_2, \dots, X_n\}$, where n is the number of observations.

Definition 1.3 (White noise). [83] A white noise process is a sequence of zero-mean, generally uncorrelated random variables $\{\varepsilon_t, t \in T\}$, where T is a set of time indices (typically $T \subseteq \mathbb{Z}$). If this sequence is Gaussian, then the process is also independent. Formally, a process $\{\varepsilon_t, t \in T\}$ is called white noise if:

- $\mathbb{E}[\varepsilon_t] = 0$ for all $t \in T$.
- $\mathrm{Var}(\varepsilon_t) = \sigma^2$ for all $t \in T$.
- $\mathrm{Cov}(\varepsilon_t, \varepsilon_s) = 0$ for all $t, s \in T$ with $t \neq s$.

If $\{\varepsilon_t, t \in T\}$ is Gaussian, then the process is called Gaussian white noise.

Definition 1.4 (Strict stationarity). [83] A stochastic process $\{X_t, t \in T\}$ is said to be strictly stationary if the joint distribution of any finite collection of random variables $\{X_{t_1}, X_{t_2}, \dots, X_{t_n}\}$ is the same as the joint distribution of $\{X_{t_1+h}, X_{t_2+h}, \dots, X_{t_n+h}\}$ for any integer h. In other words, the statistical properties of the process are invariant under time shifts.

Definition 1.5 (Weak stationarity). [83] A stochastic process $\{X_t, t \in T\}$ is said to be weakly stationary (or second-order stationary) if:

- The mean $\mathbb{E}[X_t]$ is constant over time.
- The variance $\mathrm{Var}(X_t)$ is finite and constant over time.
- The autocovariance function $\gamma(h) = \mathrm{Cov}(X_t, X_{t+h})$ depends only on the lag h and not on t.

Weak stationarity is a less restrictive condition than strict stationarity, as it only requires the first two moments (mean and variance) and the autocovariance function to be time-invariant.

Remark 1.1. A process that does not satisfy the conditions of stationarity is called non-stationary. Non-stationary processes often exhibit trends, seasonality, or other time-dependent structures.

Definition 1.6 (Autocovariance function). [83] The autocovariance function of a stochastic process $\{X_t, t \in T\}$ is defined as

$$\gamma_X(t, t') = \mathrm{Cov}(X_t, X_{t'}) = \mathbb{E}[(X_t - \mathbb{E}[X_t])(X_{t'} - \mathbb{E}[X_{t'}])],$$

where $\mathbb{E}[X_t]$ is the mean of the process.

Definition 1.7 (Autocorrelation function). [83] The autocorrelation function (ACF) of a stationary stochastic process $\{X_t, t \in T\}$ is defined as

$$\rho(h) = \frac{\gamma(h)}{\gamma(0)},$$

where

- $\gamma(h) = \mathrm{Cov}(X_{t+h}, X_t)$ is the autocovariance function at lag h,
- $\gamma(0) = \mathrm{Var}(X_t)$ is the process variance (constant under stationarity).

Proposition 1.1 (Properties of the autocorrelation function). *For a stationary stochastic process $\{X_t, t \in T\}$, the autocorrelation function satisfies the following properties:*

- $\rho(0) = 1$ *(normalization).*
- $\rho(h) = \rho(-h)$ *(symmetry).*
- $|\rho(h)| \leq 1$ *for all h (boundedness).*

Definition 1.8 (Periodically correlated processes). [31] A stochastic process $\{X_t, t \in T\}$ is said to be periodically correlated (or cyclostationary) with period $p > 0$ if its mean and covariance functions are periodic with period p, where

- The mean function satisfies $\mathbb{E}[X_{t+p}] = \mathbb{E}[X_t]$ for all t.
- The covariance function satisfies $\gamma(t + p, s + p) = \gamma(t, s)$ for all t, s.

1.2 Mathematical tools

Definition 1.9 (Lag perator ($\boldsymbol{L}$)). [83] For any time series observation X_t, the lag operator L is defined as

$$L^j X_t = X_{t-j} \quad \text{for} \quad j \in \mathbb{Z},$$

where L^j represents the lag operator applied j times, and L^0 is the identity operator ($L^0 X_t = X_t$).

Remark 1.2 (Forward operator ($\boldsymbol{F}$)). The forward operator F is the inverse of the lag operator L and is defined as

$$F^j X_t = X_{t+j} \quad \text{for} \quad j \in \mathbb{Z},$$

where F^j represents the forward operator applied j times. The inverse relationship is explicitly given by

$$L^{-1} = F \quad \text{and} \quad F^{-1} = L.$$

In particular, $L^{-1} X_t = F X_t = X_{t+1}$.

Definition 1.10 (Gamma function). [1] The gamma function, denoted by $\Gamma(z)$ where $z \in \mathbb{R}$, is defined for $\mathrm{Re}(z) > 0$ by the integral

$$\Gamma(z) = \int_0^\infty t^{z-1} e^{-t} \, dt.$$

It satisfies the following functional equations:

1. $\Gamma(z+1) = z\Gamma(z)$.

2. $\Gamma(1-z) = -z\Gamma(-z)$.

3. For $j = 1, 2, 3, \ldots$,

$$\frac{\Gamma(-z+j)}{\Gamma(-z)} = (-1)^j \frac{\Gamma(z-j+1)}{\Gamma(z)}.$$

Remark 1.3 (Gamma function's role). The gamma function is essential in extending discrete mathematical operations, such as factorials, to continuous and non-integer domains.

Definition 1.11 (Fisher information). [103] Let X be a random variable with probability density function $f(x;\theta)$, where $x \in \mathcal{X}$ and θ is an unknown parameter. The Fisher information $I(\theta)$ is defined as

$$I(\theta) = \mathbb{E}\left[\left(\frac{\partial}{\partial\theta}\log f(X;\theta)\right)^2 \middle| \theta\right].$$

If $\mathcal{X}$ does not depend on θ, this is equivalent to

$$I(\theta) = -\mathbb{E}\left[\frac{\partial^2}{\partial\theta^2}\log f(X;\theta) \middle| \theta\right].$$

The Fisher information quantifies the amount of information that X carries about the unknown parameter θ.

A larger value of $I(\theta)$ indicates greater precision in estimating θ.

1.2.1 Inequalities

1.2.1.1 Probabilistic inequalities

Theorem 1.2 (Markov inequality). *[61] Let X be a non-negative random variable, and let $a > 0$. Then,*

$$\mathbb{P}(X \geq a) \leq \frac{\mathbb{E}[X]}{a}.$$

Theorem 1.3 (Chebyshev inequality). *[61] Let X be a non-negative random variable, and let g be a positive, monotonically increasing function defined on $\mathbb{R}_+$. Then, for any $a > 0$,*

$$\mathbb{P}(X \geq a) \leq \frac{\mathbb{E}[g(X)]}{g(a)}.$$

Theorem 1.4 (Borel-Cantelli lemma). *[21]* In probability theory, the Borel-Cantelli lemma, sometimes also called the Borel-Cantelli theorem, *concerns a sequence of events. Let $\{A_n\}_{n\geq 1}$ be a sequence of events in a probability space $(\Omega, \mathcal{A}, \mathbb{P})$. If*

$$\sum_{n=1}^{\infty} \mathbb{P}(A_n) < \infty,$$

then

$$\mathbb{P}\left(\limsup_{n\to\infty} A_n\right) = 0.$$

This means that the probability of infinitely many A_n occurring is zero.

Theorem 1.5 (Prakasa-Rao inequality). *[76] Let E be a set, and let $(\delta_m(x,y))$ be a sequence of positive functions defined on $E \times E$. Let μ be a measure such that for every $\gamma > 0$,*

$$\lim_{m\to\infty} \sup_{x\in E} \left| \int_{\{y:|x-y|\leq\gamma\}} \delta_m(x,y)\mu(dy) - 1 \right| = 0.$$

Then, for any $\eta > 0$,

$$\mathbb{P}\left(|g_n(x) - \mathbb{E}[g_n(x)]| > \eta\sqrt{\frac{s_m \ln(m)}{m}} \right) \leq 2\exp\left(-\frac{s_m \ln(m)\eta^2}{8c_0 M} \right),$$

where

- $g_n(x) = \dfrac{1}{n}\sum_{i=1}^{n} \delta(x, X_i)$ *is a non-parametric estimator of the common density g of the observations X_i.*
- $\sup_{x\in E} g(x) \leq M < \infty$.
- $c_0 > 0$ *is a constant such that* $\sup_{x,y\in E} \delta_m(x,y) \leq c_0 s_m < \infty$.
- s_m *is a sequence satisfying $s_m \to \infty$ and* $\dfrac{s_m \ln(m)}{m} \to 0$ *as $m \to \infty$.*

1.2.2 General Inequalities

Theorem 1.6 (Hölder inequality). *[103] Let X and Y be random variables, and let $p, q > 1$ such that $\dfrac{1}{p} + \dfrac{1}{q} = 1$. Then,*

$$\mathbb{E}[|XY|] \leq \left(\mathbb{E}[|X|^p]\right)^{1/p} \left(\mathbb{E}[|Y|^q]\right)^{1/q}.$$

Theorem 1.7 (Cauchy-Schwarz inequality). *[103] Let X and Y be random variables. Then,*

$$\mathbb{E}[|XY|] \leq \sqrt{\mathbb{E}[X^2]\mathbb{E}[Y^2]}.$$

Theorem 1.8 (Fatou's lemma). *[61] Let $\{X_n\}_{n\in\mathbb{N}}$ be a sequence of non-negative random variables. Then,*

$$\mathbb{E}\left[\liminf_{n\to\infty} X_n\right] \leq \liminf_{n\to\infty} \mathbb{E}[X_n].$$

Theorem 1.9 (Cauchy criterion for series). *The series* $\sum_{n=1}^{\infty} a_n$ *converges if and only if for every* $\epsilon > 0$*, there exists a positive integer* N *such that for all* $m > n > N$*,*

$$\left| \sum_{j=n}^{m} a_j \right| < \epsilon.$$

1.2.3 Convergence in probability theory

1.2.3.1 Convergence Theorems

Theorem 1.10 (Monotone convergence theorem). *[61] Let* $\{X_n\}_{n\geq 0}$ *be a sequence of non-negative random variables that is monotonically increasing and converges almost surely to a random variable* X*, i.e.,*

$$X_n \to X \quad \textit{almost surely.}$$

Then,

$$\lim_{n\to\infty} \mathbb{E}[X_n] = \mathbb{E}[X].$$

Vitali's theorem [76]

Before stating the Vitali theorem, we first provide the definition of p-equi-integrability for a sequence of functions. Let the space $L^p((0,T);\chi)$ be defined as follows

$$L^p((0,T);\chi) = \left\{ u : (0,T) \to \chi \text{ measurable } ; \int_0^T \|u(t)\|_\chi^p \, dt < \infty \right\}, \quad 1 \leq p < +\infty,$$

where $(0,T)$ denotes an interval.

Definition 1.12 (p-equi-integrability). Let $1 \leq p < +\infty$. A sequence of functions $(g_n)_{n\in\mathbb{N}}$ in $L^p((0,T);\chi)$ is said to be p-equi-integrable if it satisfies the following condition: for every $\varepsilon > 0$, there exists $\delta > 0$ such that for all $n > 1$ and for every measurable subset $A \subset (0,T)$ with $|A| < \delta$, we have

$$\int_A \|g_n(t)\|_\chi^p \, dt < \varepsilon,$$

where $\|\cdot\|_\chi^p$ is a norm on χ.

Theorem 1.11 (Vitali's theorem). *Let $1 \leq p < +\infty$. If $(g_n)_{n\in\mathbb{N}}$ is a sequence in $L^p((0,T);\chi)$ that converges almost everywhere to g, then*

$$g_n \to g \text{ in } L^p((0,T);\chi)$$

if and only if $(g_n)_{n\in\mathbb{N}}$ is p-equi-integrable.

1.2.3.2 Convergence modes

Definition 1.13 (Convergence in distribution). [54] A sequence of random variables $\{X_n\}_{n\in\mathbb{N}}$ converges in distribution to a random variable X, denoted $X_n \xrightarrow{d} X$, if the sequence of cumulative distribution functions (CDF) $\{F_n\}_{n\in\mathbb{N}}$ of X_n converges pointwise to the CDF F of X at all continuity points of F,

$$\lim_{n\to\infty} F_n(x) = F(x) \quad \text{for all } x \text{ where } F \text{ is continuous.}$$

Definition 1.14 (Convergence in probability). [54] A sequence of random variables $\{X_n\}_{n\in\mathbb{N}}$ converges in probability to a random variable X, denoted $X_n \xrightarrow{p} X$, if for every $\epsilon > 0$,

$$\lim_{n\to\infty} \mathbb{P}(|X_n - X| \geq \epsilon) = 0.$$

Definition 1.15 (Almost sure convergence). [54] A sequence of random variables $\{X_n\}_{n\in\mathbb{N}}$ converges almost surely to a random variable X, denoted $X_n \xrightarrow{a.s.} X$, if

$$\mathbb{P}\left(\lim_{n\to\infty} X_n = X\right) = 1.$$

Definition 1.16 (L^p Spaces). [61] For $p \geq 1$, the L^p space is defined as

$$L^p(\Omega, \mathcal{A}, \mathbb{P}) = \{X : \Omega \to \mathbb{R} \mid X \text{ is measurable and } \|X\|_p < \infty\},$$

where the L^p-norm is given by

$$\|X\|_p = (\mathbb{E}|X|^p)^{1/p}.$$

Definition 1.17 (Convergence in L^p). [54] A sequence of random variables $\{X_n\}_{n\in\mathbb{N}}$ converges in L^p to a random variable X, denoted $X_n \xrightarrow{L^p} X$, if for $p \geq 1$,

$$\lim_{n\to\infty} \mathbb{E}\left[|X_n - X|^p\right] = 0.$$

Proposition 1.12 (Relationships between types of convergence). *The following implications hold:*

1. *almost sure convergence $\Longrightarrow$ convergence in probability $\Longrightarrow$ convergence in distribution.*

2. *convergence in L^p $\Longrightarrow$ convergence in probability $\Longrightarrow$ convergence in distribution.*

1.2.4 Slowly varying functions

Definition 1.18. [82] A function $\ell : [a, \infty) \to (0, \infty)$, where $a > 0$, is said to be slowly varying if for any constant $c > 0$,

$$\lim_{x\to\infty} \frac{\ell(cx)}{\ell(x)} = 1.$$

Examples:

- $\ell(x) = \log(x)$ is a slowly varying function.
- $\ell(x) = b$, where b is a positive constant, is also a slowly varying function.

1.2.5 Stochastic o and O symbols

In asymptotic statistics, the symbols o_P and O_P are used to describe the behavior of sequences of random variables as the sample size n tends to infinity. These notations are fundamental for characterizing stochastic convergence and boundedness. For all that follows, see [103].

Definitions

1. $o_P(1)$: A sequence of random variables X_n is $o_P(1)$ if it converges to zero in probability.i.e., for every $\epsilon > 0$,

$$\mathbb{P}(|X_n| > \epsilon) \to 0 \quad \text{as} \quad n \to \infty.$$

   This denotes terms that become negligible in probability as n grows.

2. $O_P(1)$: A sequence of random variables X_n is $O_P(1)$ if it is bounded in probability. i.e., for every $\epsilon > 0$, there exists a constant M such that

$$\mathbb{P}(|X_n| > M) < \epsilon \quad \text{for all} \quad n.$$

This denotes terms that are stochastically bounded.

3. General o_P and O_P Notation: For a sequence of random variables R_n, $X_n = o_P(R_n)$ means $X_n = Y_n R_n$ where $Y_n \xrightarrow{P} 0$. Similarly, $X_n = O_P(R_n)$ means $X_n = Y_n R_n$ where $Y_n = O_P(1)$.

Properties of o_P and O_P

The following properties hold for o_P and O_P,

$$o_P(1) + o_P(1) = o_P(1), \quad o_P(1) + O_P(1) = O_P(1), \quad O_P(1)o_P(1) = o_P(1),$$

$$(1 + o_P(1))^{-1} = O_P(1), \quad o_P(R_n) = R_n o_P(1), \quad O_P(R_n) = R_n O_P(1).$$

Lemma 1.2.1 (Stochastic $\boldsymbol{o}$ and $\boldsymbol{O}$ symbols) *Let R be a function defined on a domain in $\mathbb{R}^k$ such that $R(0) = 0$. Let X_n be a sequence of random vectors converging in probability to zero. Then, for every $p > 0$,*

1. If $R(h) = o(\|h\|^p)$ as $h \to 0$, then $R(X_n) = o_P(\|X_n\|^p)$.

2. If $R(h) = O(\|h\|^p)$ as $h \to 0$, then $R(X_n) = O_P(\|X_n\|^p)$.

This lemma is essential for analyzing the asymptotic behavior of remainder terms in stochastic expansions.

1.3 Long memory processes

Let $\gamma(h) = \langle X_t, X_{t+h} \rangle$ be the autocovariance function at lag h of the stationary process $\{X_t, t \in \mathbb{Z}\}$. A usual definition of long memory is that

$$\sum_{h=-\infty}^{\infty} |\gamma(h)| = \infty. \tag{1.1}$$

However, there are alternative definitions. In particular, long memory can be defined by specifying a hyperbolic decay of the autocovariances

$$\gamma(h) \sim h^{2d-1}\ell_1(h), \quad \text{as } h \to \infty, \tag{1.2}$$

where d is the so-called long-memory parameter and $\ell_1(\cdot)$ is a slowly varying function.

Long-memory processes include a variety of models. Among these, self-similar processes, such as fractional Brownian motion and its increments (fractional Gaussian noise), are foundational, characterized by their scaling properties and Hurst parameter. Another class includes Gegenbauer ARMA (GARMA) processes, which extend traditional ARMA models to incorporate cyclical long-memory behavior through Gegenbauer polynomials. Additionally, aggregated AR(1) processes generate long-range dependence by combining multiple autoregressive processes with heterogeneous coefficients, a method that is particularly useful in economics. Euler processes, defined through logarithmic integrals, offer another mathematical framework for modeling long-memory behavior. Finally, Fractionally Integrated ARMA (ARFIMA) processes are a highly versatile and widely used class of long-memory models. ARFIMA models are a generalization of ARMA processes, introducing fractional differencing to allow for the modeling of both short and long-range dependencies. They are a natural generalization of the ARIMA (p, d, q) models (see [40] for more details on those models' historical development). If we allow the parameter d to take any value between $-\frac{1}{2}$ and $\frac{1}{2}$, the ARFIMA process is defined by(for the sequel, we refer to [82] and [103])

$$\phi(L)(1-L)^d X_t = \theta(L)\varepsilon_t, \tag{1.3}$$

where $\phi(L) = 1 + \phi_1 L + \cdots + \phi_p L^p$ and $\theta(L) = 1 + \theta_1 L + \cdots + \theta_q L^q$ are the autoregressive and moving average operators, respectively; $\phi(L)$ and $\theta(L)$ have no common roots, L is the lag operator, d is a real number, and the polynomial $(1-L)^d$ has the expansion

$$(1-L)^d = 1 + \sum_{j=1}^{\infty} \frac{d(d-1)\cdots(d-j+1)}{j!}(-1)^j L^j. \tag{1.4}$$

$\{\varepsilon_t, t \in \mathbb{Z}\}$ is a white noise sequence with zero mean and finite variance σ^2.
Using the gamma ($\Gamma(\cdot)$) function, we can write a concise form of (1.4) as

$$(1-L)^d = \sum_{j=0}^{\infty} \frac{\Gamma(-d+j)}{\Gamma(-d)j!} L^j.$$

The ARFIMA process has the following key properties:

1. **Causality**: The process defined in (1.3) is causal if the roots of $\phi(L)$ lie outside the unit circle. This ensures:

$$X_t = \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j},$$

   where ψ_j are determined by

$$\psi_j = \frac{\theta(L)}{\phi(L)} \frac{\Gamma(j+d)}{\Gamma(d)j!}.$$

2. **Invertibility**: The process defined in (1.3) is invertible if the roots of $\theta(L)$ lie outside the unit circle. This ensures:

$$\varepsilon_t = \sum_{j=0}^{\infty} \pi_j X_{t-j},$$

   where π_j are determined by

$$\pi_j = \frac{\phi(L)}{\theta(L)} \frac{\Gamma(j-d)}{\Gamma(-d)j!}.$$

3. **Stationarity**: The process presented in (1.3) is stationary if:

   - $d < 0.5$,
   - The roots of $\phi(L)$ lie outside the unit circle (i.e., the process is causal).

   For $d \geq 0.5$, the process is non-stationary.

4. **Long-memory behavior**:

   - For $d \in (0, 0.5)$, the process exhibits long-memory, with autocorrelations decaying hyperbolically

$$\rho(h) \sim Ch^{2d-1}, \quad \text{as } h \to \infty.$$

  - For $d \in (-0.5, 0)$, the process exhibits anti-persistence, with faster decay of autocorrelations.

- When $d = 0$, the ARFIMA process reduces to an ARMA(p, q) process

$$\phi(L)X_t = \theta(L)\varepsilon_t.$$

- When d is an integer, the ARFIMA process becomes an ARIMA(p, d, q) process

$$\phi(L)(1-L)^d X_t = \theta(L)\varepsilon_t.$$

1.3.1 ARFIMA(0,d,0) processes

The ARFIMA$(0, d, 0)$ process, also denoted as $FAR(1)$ by Serroukh (1996) [94], is defined by

$$(1-L)^d X_t = \varepsilon_t, \tag{1.5}$$

The process admits a unique stationary solution for $-\frac{1}{2} < d < \frac{1}{2}$, given by

$$X_t = \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j},$$

whith

$$\psi_j = \frac{\Gamma(j+d)}{\Gamma(j+1)\Gamma(d)}.$$

Autocovariance and autocorrelation functions

The autocovariance function $\gamma_0(h)$ of the ARFIMA$(0, d, 0)$ process is

$$\gamma_0(h) = \sigma^2 \frac{\Gamma(1-2d)}{\Gamma(1-d)\Gamma(d)} \frac{\Gamma(h+d)}{\Gamma(1+h-d)},$$

and the autocorrelation function $\rho_0(h)$ is

$$\rho_0(h) = \frac{\Gamma(1-d)}{\Gamma(d)} \frac{\Gamma(h+d)}{\Gamma(1+h-d)}.$$

Classification of the process

Consider a process $\{X_t, t \in \mathbb{Z}\}$ defined by equation (1.5). For such processes, we have the following classification:

- **Stationarity**: $\{X_t, t \in \mathbb{Z}\}$ is stationary if $\sum_{j=0}^{\infty} \psi_j^2 < \infty$.
- **Memory properties**:
  - Short memory: $\sum_{j=0}^{\infty} |\psi_j| < \infty$.
  - Long memory: $\sum_{j=0}^{\infty} |\psi_j| = \infty$.
- **Non-stationarity**: $\{X_t, t \in \mathbb{Z}\}$ is non-stationary if $\sum_{j=0}^{\infty} \psi_j^2 = \infty$.

Asymptotic behavior

1. **Coefficients ψ_j**: As $j \to \infty$, the coefficients ψ_j behave as

$$\psi_j \sim \frac{j^{d-1}}{\Gamma(d)}.$$

   This hyperbolic decay is slower than exponential decay, characteristic of long memory when $0 < d < \frac{1}{2}$.

2. **Autocovariance $\gamma_0(h)$**: For large lags h, the autocovariance function $\gamma_0(h)$ satisfies

$$\gamma_0(h) \sim \sigma^2 \frac{\Gamma(1-2d)}{\Gamma(1-d)\Gamma(d)} h^{2d-1}.$$

   This shows that the autocovariance decays hyperbolically as $h \to \infty$.

3. **Memory and stationarity**:
   - For $d < \frac{1}{2}$, the process is stationary.
   - For $d \leq 0$, the process has **short memory**.
   - For $0 < d < \frac{1}{2}$, the process has **long memory**.

1.4 Estimation in long memory processes: minimum Hellinger distance estimation

Estimation in long-memory processes has been a central topic in time series analysis, with the memory parameter d historically being the primary focus of interest due to its role in characterizing long-range dependence. However, the estimation of other model parameters and the development of robust methodologies have also garnered significant attention. Classical methods for estimation in long-memory models include the maximum likelihood estimation (MLE), which provides efficient estimates under Gaussian assumptions; the Whittle estimation, a frequency-domain approach that approximates the likelihood and is computationally efficient; and the moment-based methods, such as the rescaled range (R/S) statistic (for more details and other methods, see [82]). The conditional sum of (CSS) estimator, introduced by Box and Jenkins (1970)[10], is another classical method, particularly used for ARMA and ARFIMA models. It minimizes the sum of squared conditional residuals, providing a computationally efficient alternative to MLE. However, it may be less robust in the presence of contaminated data (observations containing outliers, errors, or noise) or model deviations (situations where the true data-generating process, i.e., the underlying process that produces the data, differs from the assumed model). These methods have been widely studied and applied, offering versatile tools for analyzing long-memory processes. Although these methods are widely applicable, they are not always the most appropriate, especially when encountering non-stationary processes. In such cases, these methods may become either non-applicable or difficult to apply, and the asymptotic properties of the estimates may become challenging to study. To address these challenges, the minimum Hellinger distance estimator (MHDE) offers a robust and efficient alternative, designed to perform well even under deviations from standard assumptions.

The Hellinger distance, introduced by Ernst Hellinger in 1909, is a metric defined for two probability measures P and Q on a measurable space $(\mathcal{X}, \mathcal{A})$. It is given by

[72] as

$$HD^2(P,Q) = \frac{1}{2}\int_{\mathcal{X}} \left(\sqrt{dP} - \sqrt{dQ}\right)^2,$$

where $\sqrt{dP}$ and $\sqrt{dQ}$ are the square roots of the densities of P and Q with respect to a dominating measure μ (for example, the Lebesgue measure for continuous distributions).

The MHDE minimizes the Hellinger distance between the empirical distribution P_n of the observed data and the parametric model distribution P_θ. Formally, given a parametric family $\{P_\theta : \theta \in \Theta\}$, where Θ is the parameter space, the MHDE is defined as

$$\hat{\theta}_{\text{MHDE}} = \arg\min_{\theta\in\Theta} HD^2(P_n, P_\theta).$$

This estimator seeks to find the parameter θ that makes the model distribution P_θ as close as possible to the empirical distribution P_n in terms of the Hellinger distance. Historically, the Hellinger distance gained prominence in statistics through the work of Kakutani (1948) [58], who used it to study the absolute continuity of product measures. Later, Kraft (1955) [65] and Matusita (1955) [74] further developed its applications in statistical inference. The MHDE was formally introduced by Beran (1977) [8] for parametric models, who established its robustness properties. Specifically, Beran showed that the MHDE is asymptotically efficient under the true model, achieving the Cramér-Rao lower bound, and remains consistent even under slight deviations from the assumed model, making it a robust alternative to MLE. The asymptotic properties of MHDE, including consistency and asymptotic normality, further improve its applicability to ARFIMA models, where the slow decay of correlations can complicate the behavior of estimators. Later, Tamura and Boos (1986) [99] extended the MHDE to multivariate processes, enabling its application to higher-dimensional and more structured statistical frameworks, such as multivariate distributions, regression, and time series models, while preserving its robustness and efficiency properties. More recently, Mbeke (2019) [76] constructed MHDE for several multivariate stationary long-memory processes (see [72] for more historical background and [110] for the mathematical development of MHDE for parametric models, multivariate models, and others).

1.5 Local asymptotic normality

Local asymptotic normality (LAN) describes how parametric statistical models behave in small neighborhoods around the true parameter value θ_0 as the sample size n increases. The sequence of models $\{P_{\theta,n}\}_{\theta\in\Theta}$ has the LAN property if the log-likelihood ratio process can be approximated quadratically

$$\log\left(\frac{dP_{\theta_0+h/\sqrt{n},n}}{dP_{\theta_0,n}}\right) = h^\top \Delta_n - \frac{1}{2}h^\top I_{\theta_0} h + o_P(1),$$

where $\Delta_n \rightsquigarrow N(0, I_{\theta_0})$ and I_{θ_0} is the Fisher information matrix. This means that, locally, the statistical model behaves like a Gaussian shift model $N(h, I_{\theta_0}^{-1})$.
A standard example is when we observe n i.i.d. samples from a regular parametric model (such as normal distributions), where the usual $\sqrt{n}$-scaling leads to the LAN approximation. The following subsections rigorously develop these ideas, building on the foundational contributions of [72] and [103].

1.5.1 Statistical experiment framework

Consider a statistical experiment $\mathcal{E} = \{P_\theta : \theta \in \Theta \subset \mathbb{R}^k\}$ defined on a sample space $(\mathcal{X}, \mathcal{A})$. Given n i.i.d. observations, the joint model is the product experiment $\mathcal{E}_n = \{P_\theta^n : \theta \in \Theta\}$.
Under regularity conditions, $\mathcal{E}_n$ admits a local asymptotic normal approximation around a fixed $\theta_0 \in \Theta$ via the reparametrization

$$h = \sqrt{n}(\theta - \theta_0), \quad \text{for } h \in \mathbb{R}^k.$$

This yields the localized experiment $\widetilde{\mathcal{E}}_n = \{P^n_{\theta_0+h/\sqrt{n}} : h \in \mathbb{R}^k\}$.

1.5.2 LAN expansion

Assume the densities p_θ satisfy the following:

1. $\log p_\theta(x)$ is twice differentiable in θ.

2. The Fisher information matrix $I_{\theta_0} = -\mathbb{E}_{\theta_0}\left[\frac{\partial^2 \log p_\theta}{\partial\theta\partial\theta^\top}\bigg|_{\theta=\theta_0}\right]$ is positive definite.

Then, the log-likelihood ratio process admits the quadratic approximation

$$\log \frac{dP^n_{\theta_0+h/\sqrt{n}}}{dP^n_{\theta_0}} = h^\top \Delta_n(\theta_0) - \frac{1}{2}h^\top I_{\theta_0} h + o_{P^n_{\theta_0}}(1), \tag{1.6}$$

where

- $\Delta_n(\theta_0) = \frac{1}{\sqrt{n}}\sum_{i=1}^n \frac{\partial \log p_\theta}{\partial\theta}\bigg|_{\theta=\theta_0}(X_i)$ is th score vector, satisfying $\Delta_n(\theta_0) \xrightarrow{\mathcal{L}} \mathcal{N}(0, I_{\theta_0})$ under $P^n_{\theta_0}$ by the central limit theorem.
- The remainder term $o_{P^n_{\theta_0}}(1)$ converges to zero in probability.

1.5.3 Gaussian approximation

The LAN expansion (1.6) implies that $\widetilde{\mathcal{E}}_n$ is asymptotically equivalent to the Gaussian shift experiment

$$\mathcal{G} = \{\mathcal{N}(h, I^{-1}_{\theta_0}) : h \in \mathbb{R}^k\},$$

where the log-likelihood ratio for $\mathcal{G}$ is

$$\log \frac{d\mathcal{N}(h, I^{-1}_{\theta_0})}{d\mathcal{N}(0, I^{-1}_{\theta_0})}(X) = h^\top I_{\theta_0} X - \frac{1}{2}h^\top I_{\theta_0} h.$$

This equivalence follows from matching:

- The linear term $h^\top \Delta_n(\theta_0)$ corresponds to $h^\top I_{\theta_0} X$.
- The quadratic term $-\frac{1}{2}h^\top I_{\theta_0} h$ is identical in both models.

1.5.4 Implications for inference

Theorem 1.13 (Asymptotic efficiency under LAN). *Under LAN conditions:*

1. *The MLE $\hat{\theta}_n$ satisfies $\sqrt{n}(\hat{\theta}_n - \theta_0) \xrightarrow{\mathcal{L}} \mathcal{N}(0, I^{-1}_{\theta_0})$.*
2. *Any regular estimator sequence has a limiting distribution that is a convolution of $\mathcal{N}(0, I^{-1}_{\theta_0})$ with another distribution (Hájek-Le Cam convolution theorem).*

CHAPTER 2

FRACTIONAL PERIODIC AUTOREGRESSION

2.1 Introduction

The present chapter highlights the initial probabilistic properties of a new class of fractional autoregressive models, developed based on the model given in 1987 by Gonçalves [33]. We introduce the periodicity in the structure of the stochastic equation of the classical model, providing it with the periodicity property. We give it causality and invertibility conditions; then we show that the process that satisfies the model is non-stationary and prove that it can be defined as periodically correlated. We also present the periodic autocovariance and autocorrelation functions with their asymptotic forms and behavior.

We give a graphical study at the end to visualize the main results and extract the effects of introducing the periodicity.

2.2 Fractional process with periodic coefficient

2.2.1 Definition and notations

A stochastic process $\{X_t, t \in \mathbb{Z}\}$ is said to be the fractional process with a periodic coefficient if it has the following representation

$$(1 - a_t L)^d X_t = \varepsilon_t, \tag{2.1}$$

where

- L is a lag operator ($L^j X_t = X_{t-j}$).
- (a_t, d) are unknown parameters; d is not necessarily an integer.
- The innovation process $\{\varepsilon_t, t \in \mathbb{Z}\}$ denotes a sequence of independent and identically distributed random variables with zero mean and finit variance σ^2.

We consider the series of fractional operator $A_t(L) = (1 - a_t L)^d$. Using the binomial series, we develop it formula as follows

$$\begin{aligned}
A_t(L) &= (1 - a_t L)^d \\
&= \sum_{j=0}^{\infty} \binom{d}{j} 1^{d-j} (-a_t L)^j \\
&= \sum_{j=0}^{\infty} \frac{d!}{(d-j)!j!} (-1)^j (a_t L)^j \\
&= \sum_{j=0}^{\infty} \frac{d(d-1)...(d-j+1)(d-j)!}{(d-j)!j!} (-1)^j (a_t L)^j \\
&= \sum_{j=0}^{\infty} \frac{(j-d-1)...(1-d)(-d)}{j!} (a_t L)^j,
\end{aligned}$$

when we introduce the gamma function ($\Gamma(.)$), we get to the following form

$$A_t(L) = \sum_{j=0}^{\infty} \frac{\Gamma(j-d)}{\Gamma(-d)j!} a_t^j L^j.$$

In this work a_t is assumed to be periodic with period $s \in \mathbb{N}$; thus, we have the following:

for all $t \in \mathbb{Z}$, $\exists i \in \{1, ..., s\}$ and $M \in \mathbb{Z}$ such that $t = i + sM$ and so $a_t = a_{i+sM} = a_i$, from which it results

$$A_t(L) = \sum_{j=0}^{\infty} \frac{\Gamma(j-d)}{\Gamma(-d)j!} a_i^j L^j = \sum_{j=0}^{\infty} \Pi_j(i) L^j,$$

with

$$\Pi_j(i) = \frac{\Gamma(j-d)}{\Gamma(-d)\Gamma(j+1)} a_i^j. \tag{2.2}$$

Same as $A_t(L)$, we reformulate its inverse operator given by $A_t'(L) = (1 - a_t L)^{-d}$ and we get

$$A_t'(L) = \sum_{j=0}^{\infty} \frac{\Gamma(j+d)}{\Gamma(d)j!} a_i^j L^j = \sum_{j=0}^{\infty} \Psi_j(i) L^j,$$

with

$$\Psi_j(i) = \frac{\Gamma(j+d)}{\Gamma(d)\Gamma(j+1)} a_i^j. \tag{2.3}$$

(I) The process defined by the relation (2.1) has an infinite moving average representation ($MA(\infty)$) if $|a_i| < 1$ or $a_i = \pm 1$ and $d < \frac{1}{2}$ as follows

$$X_{i+sM} = \sum_{j=0}^{\infty} \Psi_j(i) \varepsilon_{i+sM-j}. \tag{2.4}$$

(II) The process defined by the relation (2.1) is invertible, and it can be written under infinite autoregressive representation ($AR(\infty)$) if $|a_i| < 1$ or $a_i = \pm 1$ and $d > -\frac{1}{2}$ as follows

$$\varepsilon_{i+sM} = \sum_{j=0}^{\infty} \Pi_j(i) X_{i+sM-j}. \tag{2.5}$$

We can now write (2.1) in the following form

$$(1 - a_i L)^d X_{i+sM} = \varepsilon_{i+sM}. \tag{2.6}$$

2.2.2 Causality and invertibility condition

Considering the model given in (2.6), from the univariate process $\{X_t, t \in \mathbb{Z}\}$, presented by (2.4), we can form a multivariate process $\{Y_M, M \in \mathbb{Z}\}$ by forming $s-$dimensional vectors of s consecutive $X's$ as follows

$$Y_M = \sum_{j=0}^{\infty} \mathit{\Psi}_j \eta_{M-j}, \tag{2.7}$$

such that

$$\eta_M = \begin{pmatrix} \varepsilon_{1+sM} \\ \varepsilon_{2+sM} \\ . \\ . \\ \varepsilon_{s+sM} \end{pmatrix}, \qquad Y_M = \begin{pmatrix} X_{1+sM} \\ X_{2+sM} \\ . \\ . \\ X_{s+sM} \end{pmatrix}, \qquad \mathit{\Psi}_j = \begin{pmatrix} \Psi_j(1) & 0 & \cdots & 0 \\ 0 & \Psi_j(2) & 0 & \vdots \\ \vdots & 0 & \ddots & 0 \\ 0 & \vdots & 0 & \Psi_j(s) \end{pmatrix},$$

this transition to a vectorial form and the results of Serroukh [94] enable us to consider a sufficient condition of causality and invertibility.
From Serroukh [94], we distinguish the behavior of the moving average coefficients $(\Psi_j(i))$ and the autoregressive coefficients $(\Pi_j(i))$ according to the values of a_i and d.
If $a_i = a =$ constant, then

- For $|a| < 1$, $\Psi_j(i)$ and $\Pi_j(i)$ are summable.
- For $a = \pm 1$
  - if $d < 0$ $\Psi_j(i)$ then are summable.
  - if $0 < d < \frac{1}{2}$ then $\Psi_j(i)$ are not summable but square summable.
  - if $d \geq \frac{1}{2}$ then $\Psi_j(i)$ are neither summable nor square summable.
  - if $d > 0$ then $\Pi_j(i)$ are summable.
  - if $-\frac{1}{2} < d < 0$ then $\Pi_j(i)$ are not summable but square summable.
  - if $d \leq -\frac{1}{2}$ then $\Pi_j(i)$ are neither summable nor square summable.

So the sufficient condition of causality and invertibility is that $\forall i = \{1, ..., s\}$, $|a_i| < 1$ or $a_i = \pm 1$ and $-\frac{1}{2} < d < \frac{1}{2}$.

Remark 2.1. For the remainder of the paper, we suppose that for all $i = \{1, ..., s\}$, $|a_i| < 1$ and $d \in \mathbb{R}$.

Proposition 2.1. *Let $\{X_t, t \in \mathbb{Z}\}$ be a process satisfying the stochastic equation (2.6); $\forall i = \{1, ..., s\}$ if $|a_i| < 1$, each component of $\{Y_M, M \in \mathbb{Z}\}$ is causal and* *invertible, and the infinite series* $\sum_{j\geq 0} \Psi_j(i)\varepsilon_{i+sM-j}$ *and* $\sum_{j\geq 0} \Pi_j(i)X_{i+sM-j}$ *converge in root mean square.*

Proof 1 *For all* $M \in \mathbb{Z}, i = \{1, ..., s\}$*, the coefficients (2.3) and let* T *and* R *be two positive integers such that* $T < R$*, we define* $S_T = \sum_{j=0}^{T} \Psi_j(i)\varepsilon_{i+sM-j}$ *and we calculate* $\mathbb{E}[|S_R - S_T|^2]$ *as follows*

$$\begin{aligned}
\mathbb{E}[|S_s - S_t|^2] &= \mathbb{E}[|\sum_{j=0}^{R} \Psi_j(i)\varepsilon_{i+sM-j} - \sum_{j=0}^{T} \Psi_j(i)\varepsilon_{i+sM-j}|^2] \\
&= \mathbb{E}[\sum_{j=T+1}^{R} (\Psi_j(i))^2\varepsilon^2_{i+sM-j} + \sum_{k,j=T+1,k\neq j}^{R} \Psi_j(i)\Psi_j(i)\varepsilon_{i+sM-k}\varepsilon_{i+sM-j}] \\
&= \sigma^2 \sum_{j=T+1}^{R} (\Psi_j(i))^2,
\end{aligned}$$

if we write $j = l + sM$*,* $l = \{1, ..., s\}$*, with easy simplifications, we get*

$$\mathbb{E}[|S_R - S_T|^2] = \sigma^2 \sum_{l=1}^{s} \sum_{M=[\frac{T}{s}]+1}^{[\frac{R}{s}]} (\Psi_{l+sM}(i))^2,$$

by Cauchy's criteria, it is easy to show that

$\sigma^2 \sum_{l=1}^{s} \sum_{M=[\frac{T}{s}]+1}^{[\frac{R}{s}]} (\Psi_{l+sM}(i))^2 < \infty$ *for all* $i = \{1, ..., s\}$*, hence the series* $\sum_{j\geq 0} \Psi_j(i)\varepsilon_{i+sM-j}$ *converges in root mean square.*

Proving the convergence of $\sum_{j\geq 0} \Pi_j(i)X_{i+sM-j}$ *is analogous.*

2.3 Autocovariance and autocorrelation functions

Let $\{X_t, t \in \mathbb{Z}\}$ be a process satisfying the stochastic equation (2.6); the covariance function of $\{X_t, t \in \mathbb{Z}\}$ is defined by

$$\begin{aligned}
cov(X_t, X_{t+h}) &= \mathbb{E}\{(X_t - \mathbb{E}(X_t))(X_{t+h} - \mathbb{E}(X_{t+h}))\} \\
&= \mathbb{E}\left\{(\sum_{j\geq 0} \Psi_j(t)\varepsilon_{t-j})(\sum_{j\geq 0} \Psi_j(t+h)\varepsilon_{t+h-j})\right\} \\
&= \sum_{j\geq 0} \Psi_j(t)\Psi_j(t+h)\mathbb{E}(\varepsilon^2_{t-j}) \\
&= \sigma^2 \sum_{j\geq 0} \Psi_j(t)\Psi_j(t+h).
\end{aligned}$$

We can see that there is a t figuring in the final form of the covariance function. This signifies that the process $\{X_t, t \in \mathbb{Z}\}$ is non-stationary (when $a_t = a$ it is simple to check that $\{X_t, t \in \mathbb{Z}\}$ is stationary).

A specific category of non-stationary processes is characterized by periodic autocovariance functions. Introduced by Gladyshev in 1961 [31], these processes are formally known as periodically correlated processes or cyclostationary processes. They exhibit structured non-stationarity, where statistical properties such as the mean and autocovariance vary periodically over time. This remark leads to the following proposition.

Proposition 2.2 (periodically correlated process). *Let $\{X_t, t \in \mathbb{Z}\}$ be a process satisfying the stochastic equation (2.6), $\{X_t, t \in \mathbb{Z}\}$ is a periodically correlated process.*

Proof 2 *The proof of the proposition 2.2 is grounded on two points:*
(i) From the infinite moving average representation (2.4), it is clear that

$$\mathbb{E}(X_{t+sM}) = 0 = \mathbb{E}(X_t).$$

(ii) We have

$$\Psi_j(t) = \frac{\Gamma(j+d)}{\Gamma(d)\Gamma(j+1)} a_t = \frac{\Gamma(j+d)}{\Gamma(d)\Gamma(j+1)} a_{t+sM} = \Psi_j(t+sM),$$

we denote by $\gamma_x(t)$ the autocovariance function, and we have

$$\begin{aligned}
\gamma_x(t+sM, t+sM+h) &= \mathbb{E}\left[(X_{t+sM} - \mathbb{E}(X_{t+sM}))(X_{t+sM+h} - \mathbb{E}(X_{t+sM+h}))\right] \\
&= \mathbb{E}[(\sum_{j\geq 0} \Psi_j(t+sM)\varepsilon_{t+sM-j})(\sum_{j\geq 0} \Psi_j(t+sM+h)\varepsilon_{t+sM+h-j})] \\
&= \sigma^2 \sum_{j\geq 0} \Psi_j(t+sM)\Psi_j(t+sM+h) \\
&= \sigma^2 \sum_{j\geq 0} \Psi_j(t)\Psi_j(t+h) \\
&= \gamma_x(t, t+h).
\end{aligned}$$

Both $\mathbb{E}(X_t)$ and $\gamma_x(t)$ are periodic with the same period of a_t from (i) and (ii) with $|a_t| < 1$, so we say that $\{X_t, t \in \mathbb{Z}\}$ is a periodically correlated process.

Notation

We set $\gamma_x(t, t+h) = \gamma_x^i(h)$ where $t = i + sM$.

Remark 2.2. (2.7) is the stationary form of the process $\{X_t, t \in \mathbb{Z}\}$.

Proposition 2.3. *The autocovariance and autocorrelation (denoted by $\rho_x^i(h)$) functions of the process $\{X_t, t \in \mathbb{Z}\}$ satisfy*

$$\gamma_x^i(h) = \sigma^2 a_{i+h}^h \frac{\Gamma(d+h)}{\Gamma(d)\Gamma(h+1)} F(d, d+h, h+1, a_i a_{i+h}),$$

$$\rho_x^i(h) = a_{i+h}^h \frac{\Gamma(d+h) F(d, d+h, h+1, a_i a_{i+h})}{\Gamma(d)\Gamma(h+1) F(d, d, 1, a_i^2)},$$

when h tends to $+\infty$, we have

$$\gamma_x^i(h) \sim \sigma^2 a_{i+h}^h \frac{h^{d-1}}{\Gamma(d)(1 - a_i a_{i+h})^d},$$

$$\rho_x^i(h) \sim a_{i+h}^h \frac{h^{d-1}}{\Gamma(d)(1 - a_i a_{i+h})^d F(d, d, 1, a_i^2)},$$

where F is the hypergeometric function.

Proof 3 *We found in the proof of the proposition 2.2 that the autocovariance function can be written as follows*

$$\begin{aligned}
\gamma_x^i(h) &= \sigma^2 \sum_{j \geq 0} \Psi_j(i) \Psi_{j+h}(i+h) \\
&= \sigma^2 \sum_{j \geq 0} \frac{\Gamma(j+d)\Gamma(j+d+h)}{\Gamma(d)\Gamma(j+1)\Gamma(d)\Gamma(j+h+1)} a_i^j a_{i+h}^{j+h}.
\end{aligned}$$

- *(i) The expression $\gamma_x^i(h)$ is obtained immediately from the form provided earlier, after having used the appropriate properties of the hypergeometric function [1] such that*

$$F(d, d+h, h+1, a_i a_{i+h}) = \frac{\Gamma(h+1)}{\Gamma(d)\Gamma(d+h)} \sum_{j \geq 0} \frac{\Gamma(d+j)\Gamma(d+h+j)}{\Gamma(h+1+j)\Gamma(j+1)} (a_i a_{i+h})^j.$$

- *(ii) The autocorrelation function is found by dividing $\gamma_x^i(h)$ by $\gamma_x^i(0)$ such that*

$$\gamma_x^i(0) = Var(X_t) = \sigma^2 F(d, d, 1, a_i^2).$$

- *(iii) For the approximations of $\gamma_x^i(h)$ and $\rho_x^i(h)$ (when h tends to $+\infty$), we consider the Schepard formula to replace $\dfrac{\Gamma(h+d)}{\Gamma(h+1)}$ and $F(d, d+h, h+1, a_i a_{i+h})$ by their approximate values [34] defined by*

$$\frac{\Gamma(h+d)}{\Gamma(h+1)} \sim h^{d-1},$$

and

$$F(d, d+h, h+1, a_i a_{i+h}) \sim \frac{1}{(1-a_i a_{i+h})^d}.$$

Remark 2.3. In the previous proposition, we see that $\gamma_x^i(h)$ is derived in terms of a_i, $i = \{1, ..., s\}$, so in order to clarify the effect of the periodicity on it, we must specify its formula in terms of various values of s where $s \in \mathbb{N}$.

2.3.1 Periodic fractional autoregressive model with period $s \in \mathbb{N}$

In this subsection, we detail the results of proposition 2.3 for various values of a_i, $i = \{1, ..., s\}$. We begin with a general value of s where $s \in \mathbb{N}$.

Proposition 2.4. *Let $s \in \mathbb{N}$, for each $i \in \{1, ..., s\}$ and $k \in \mathbb{N}$ the autocovariance functions of the model (2.6) are given by*

$$\gamma_x^i(h) = \begin{cases} \sigma^2 a_{i+1}^h \dfrac{\Gamma(d+h)}{\Gamma(d)\Gamma(h+1)} F(d, d+h, h+1, a_i a_{i+1}), & if \quad h \equiv 1[s], \\ \vdots & \\ \sigma^2 a_{i+k}^h \dfrac{\Gamma(d+h)}{\Gamma(d)\Gamma(h+1)} F(d, d+h, h+1, a_i a_{i+k}), & if \quad h \equiv k[s], \\ \vdots & \\ \sigma^2 a_{i+s}^h \dfrac{\Gamma(d+h)}{\Gamma(d)\Gamma(h+1)} F(d, d+h, h+1, a_i a_{i+s}), & if \quad h \equiv s[s], \end{cases}$$

Furthermore, the autocorrelation functions are given by

$$\rho_x^i(h) = \begin{cases} a_{i+1}^h \dfrac{\Gamma(d+h)}{\Gamma(d)\Gamma(h+1)} \dfrac{F(d, d+h, h+1, a_i a_{i+1})}{F(d, d, 1, a_i^2)}, & if \quad h \equiv 1[s], \\ \vdots & \\ a_{i+k}^h \dfrac{\Gamma(d+h)}{\Gamma(d)\Gamma(h+1)} \dfrac{F(d, d+h, h+1, a_i a_{i+k})}{F(d, d, 1, a_i^2)}, & if \quad h \equiv k[s], \\ \vdots & \\ a_{i+s}^h \dfrac{\Gamma(d+h)}{\Gamma(d)\Gamma(h+1)} \dfrac{F(d, d+h, h+1, a_i a_{i+s})}{F(d, d, 1, a_i^2)}, & if \quad h \equiv s[s]. \end{cases}$$

Proof 4 *Same as (i) and (ii) in the proof of the proposition 2.3.*

Corollary 2.1. *Using the standard approximation derived from Stirling's formula for large h and according to the proposition above, the approximations of the functions are given by.*

$$\gamma_x^i(h) \sim \begin{cases} \sigma^2 a_{i+1}^h \dfrac{h^{d-1}}{\Gamma(d)(1 - a_i a_{i+1})^d}, & if \quad h \equiv 1[s], \\ \vdots & \\ \sigma^2 a_{i+k}^h \dfrac{h^{d-1}}{\Gamma(d)(1 - a_i a_{i+k})^d}, & if \quad h \equiv k[s], \\ \vdots & \\ \sigma^2 a_{i+s}^h \dfrac{h^{d-1}}{\Gamma(d)(1 - a_i a_{i+s})^d}, & if \quad h \equiv s[s], \end{cases}$$

and

$$\rho_x^i(h) \sim \begin{cases} a_{i+1}^h \dfrac{h^{d-1}}{\Gamma(d)(1 - a_i a_{i+1})^d F(d, d, 1, a_i^2)}, & if \quad h \equiv 1[s], \\ \vdots & \\ a_{i+k}^h \dfrac{h^{d-1}}{\Gamma(d)(1 - a_i a_{i+k})^d F(d, d, 1, a_i^2)}, & if \quad h \equiv k[s], \\ \vdots & \\ a_{i+s}^h \dfrac{h^{d-1}}{\Gamma(d)(1 - a_i a_{i+s})^d F(d, d, 1, a_i^2)}, & if \quad h \equiv s[s]. \end{cases}$$

Proof 5 *Analogous to (iii) of the proof of the proposition 2.3.*

2.3.2 Periodic fractional autoregressive model with period s=2

This subsection deals with a particular case where we present the results for a period $s = 2$, i.e., $i = 1$ or $i = 2$.

Corollary 2.2. *For $s = 2$ then $i = 1$ or 2 the autocovariance functions of the model (2.6) are given by*

$$\gamma_x^i(h) = \sigma^2 \begin{cases} a_i^h \dfrac{\Gamma(d+h)}{\Gamma(d)\Gamma(h+1)} F(d, d+h, h+1, a_i^2), & if \quad h \quad even, \\ a_{i+1}^h \dfrac{\Gamma(d+h)}{\Gamma(d)\Gamma(h+1)} F(d, d+h, h+1, a_i a_{i+1}), & if \quad h \quad odd, \end{cases}$$

also the autocorrelation functions are represented by

$$\rho_x^i(h) = \begin{cases} a_i^h \dfrac{\Gamma(d+h)}{\Gamma(d)\Gamma(h+1)} \dfrac{F(d, d+h, h+1, a_i^2)}{F(d, d, 1, a_i^2)}, & if \quad h \quad even, \\ a_{i+1}^h \dfrac{\Gamma(d+h)}{\Gamma(d)\Gamma(h+1)} \dfrac{F(d, d+h, h+1, a_i a_{i+1})}{F(d, d, 1, a_i^2)}, & if \quad h \quad odd. \end{cases}$$

Proof 6 *This is similar to the proof of the proposition 2.4.*

Corollary 2.3. *According to the corollary 2.2, when $h \to +\infty$ we have*

$$\gamma_x^i(h) \sim \begin{cases} \sigma^2 \dfrac{a_i^h h^{d-1}}{\Gamma(d)(1-a_i^2)^d}, & if \quad h \quad even, \\ \sigma^2 \dfrac{a_{i+1}^h h^{d-1}}{\Gamma(d)(1-a_i a_{i+1})^d}, & if \quad h \quad odd, \end{cases}$$

in this limiting case, $(h \longrightarrow \infty)$, the autocovariance function shows a sort of hyperbolic decay; furthermore, we have

$$\rho_x^i(h) \sim \begin{cases} \dfrac{a_i^h h^{d-1}}{\Gamma(d)(1-a_i^2)^d F(d, d, 1, a_i^2)}, & if \quad h \quad even, \\ \dfrac{a_{i+1}^h h^{d-1}}{\Gamma(d)(1-a_i a_{i+1})^d F(d, d, 1, a_i a_{i+1})}, & if \quad h \quad odd, \end{cases}$$

Proof 7 *In one hand to deduce the asymptotic behavior of the autocovariance functions, we consider Sheppard formula, we have*

$$\frac{\Gamma(h+d)}{\Gamma(h+1)} \sim h^{d-1},$$

and

$$F(d, d+h, h+1, a_i a_{i+1}) \sim \frac{1}{(1-a_i a_{i+1})^d},$$

on the other hand, it is known that $\rho_x^i(h) = \dfrac{\gamma_x^i(h)}{\gamma_x^i(0)}$, *so by dividing each of the autocovariances by*

$$\gamma_x^i(0) = \begin{cases} \sigma^2 F(d, d, 1, a_i^2), & if \quad h \quad even, \\ \sigma^2 F(d, d, 1, a_i a_{i+1}), & if \quad h \quad odd, \end{cases}$$

we find the given result.

Proposition 2.5. *The autocovariance function presented in the proposition 2.3 satisfies the following recurrence relation for every* $h \in \mathbb{Z}$, $i = \{1, ..., s\}$ *and for* $(i+h) \equiv k[s]$

$$\frac{1}{\sigma^2 a_{k+2}^{h+2}} \frac{\Gamma(d)\Gamma(h-1)}{\Gamma(d+h-2)} \left[\frac{h}{a_i a_k + 1}\right] \gamma_x^i(h-2) + \frac{1}{\sigma^2 a_k^h} \frac{\Gamma(d)\Gamma(h-1)}{\Gamma(d+h)} \times$$

$$\left[\frac{(d+h-1)(h-d)(a_i a_k)}{(h-1)a_i a_k + (h-1)} - h(a_i a_k + 1) + \frac{(d+h)a_i a_k}{a_i a_k + 1}\right] \gamma_x^i(h) +$$

$$\frac{1}{\sigma^2 a_{k+2}^{h+2}} \frac{\Gamma(d)\Gamma(h+3)}{\Gamma(d+h+2)} \left[\frac{(h+1-d)(h+d+1)(h-d+2)(a_i a_k)^2}{-(h+2)(a_i a_k + (h+1))}\right] \gamma_x^i(h+2) = 0.$$

Proof 8 *Let the following expression, which is derived using Gauss contiguous relations:*

$$(c-1)F(a, b-1, c-1, z) + (az - bz - c + 1)F(a, b, c, z) + \frac{b(c-a)}{c} F(a, b+1, c+1, z) = 0. \tag{2.8}$$

In (2.8), if we replace $F(a, b-1, c-1, z)$ *and* $F(a, b+1, c+1, z)$ *by their forms such that*

$$F(a, b-1, c-1, z) = \frac{-1}{(az - (b-1)z - (c-1) + 1}(c-2)F(a, b-2, c-2, z) + \frac{(b-1)(c-1-a)}{c-1} zF(a, b, c, z), \quad \text{and,}$$

$$\begin{aligned}F(a,b+1,c+1,z) =& \frac{-1}{(az-(b+1)z-(c+1)+1}cF(a,b,c,z)+\\ & \frac{(b+1)(c+1-a)}{c+1}zF(a,b+2,c+2,z),\end{aligned}$$

after some simplifications, we get to

$$\begin{aligned}&F(a,b-2,c-2,z)\left[\frac{-(c-1)(c-2)}{az-(b-1)z-c+2}+F(a,b,c,z)\right]\\ &\left[\frac{-(b-1)(c-1-a)z}{az-(b-1)z-c+2}+(az-bz-c+1)+\frac{-b(c-a)z}{az-(b+1)z-c}\right]+\\ &F(a,b+2,c+2,z)\left[\frac{-b(c-a)(b+1)(c+1+a)}{c(c+1)(az-(b+1)z-c)}z^2\right]=0.\end{aligned} \tag{2.9}$$

We set the following $a=d$*,* $b=d+h$*,* $c=h+1$ *and* $z=a_ia_{i+h}$*.*

On an other hand we have

$$\gamma_x^i(h)=\sigma^2a_k^h\frac{\Gamma(d+h)}{\Gamma(d)\Gamma(h+1)}F(d,d+h,h+1,a_ia_k),\quad if\quad i+h\equiv k[s],$$

from which we write $F(d,d+h,h+1,a_ia_k)=\gamma_x^i(h)\dfrac{1}{\sigma^2a_k^h}\dfrac{\Gamma(d)\Gamma(h+1)}{\Gamma(d+h)}$.

When we substitute each hypergeometric function in (2.9) by its expression such that

$$F(a,b-2,c-2,z)=F(d,(d+h)-2,(h+1)-2,a_ia_k),\quad \text{and,}$$

$$F(a,b+2,c+2,z)=F(d,(d+h)+2,(h+1)+2,a_ia_k),$$

after some calculations, we find the result cited in the proposition 2.5.

2.4 Simulation

This section is devoted to a simulation view of the periodic form of the autocovariance functions.

We provide multiple figures for $\gamma_x^i(h)$ to examine the effects of the periodicity on the process. For this purpose, we consider two different periods.

First, we take $s=2$, where its theoretical results are presented in the corollary 2.2; then we highlight the results of the proposition 2.4 and consider a period $s=4$. For each case, we examine several values of the parameters a_i, $i=\{1,...,s\}$ and d.

In all the following figures we consider $h=100$.

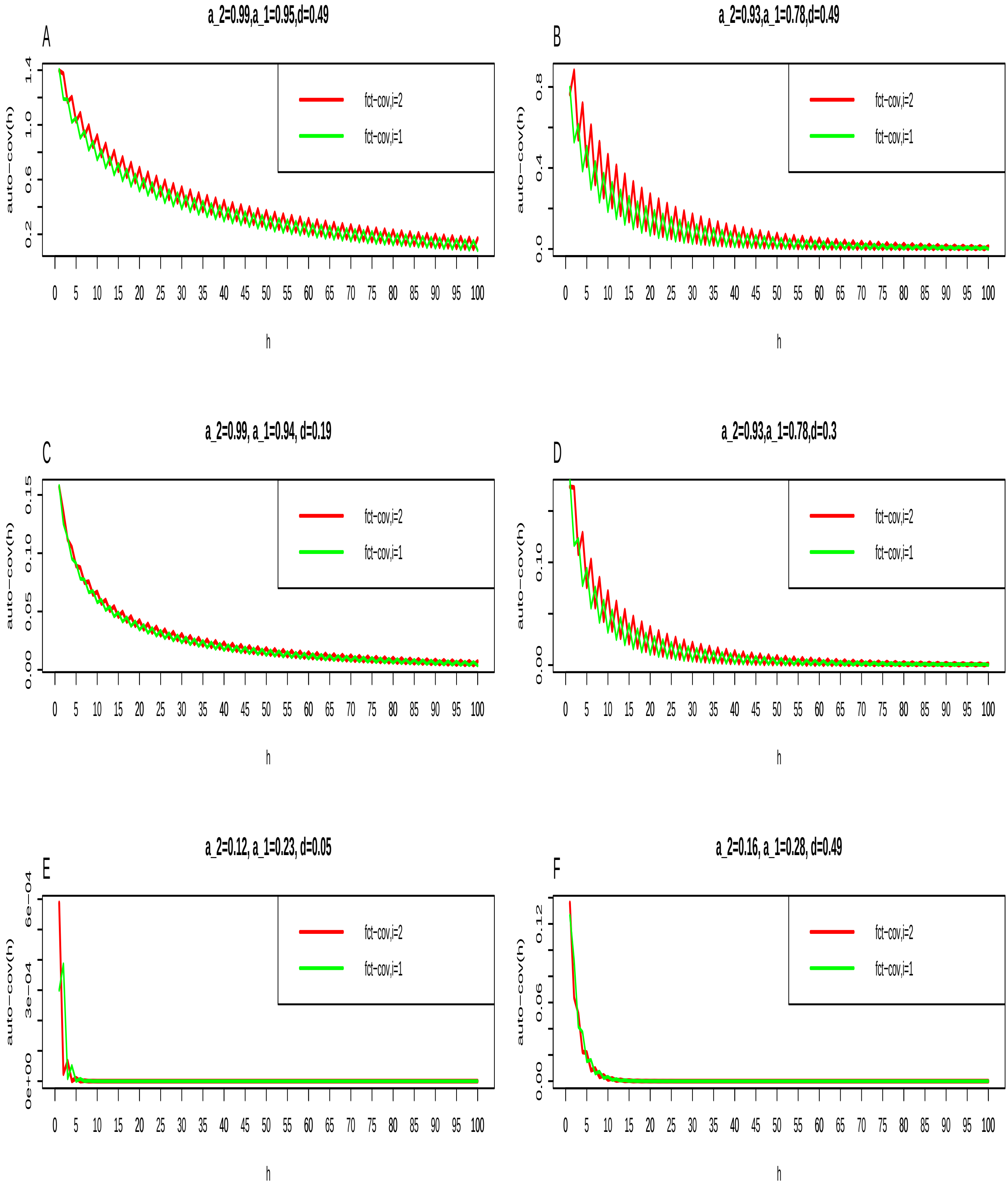

Figure 2.1: Periodic autocovariance $s = 2$.

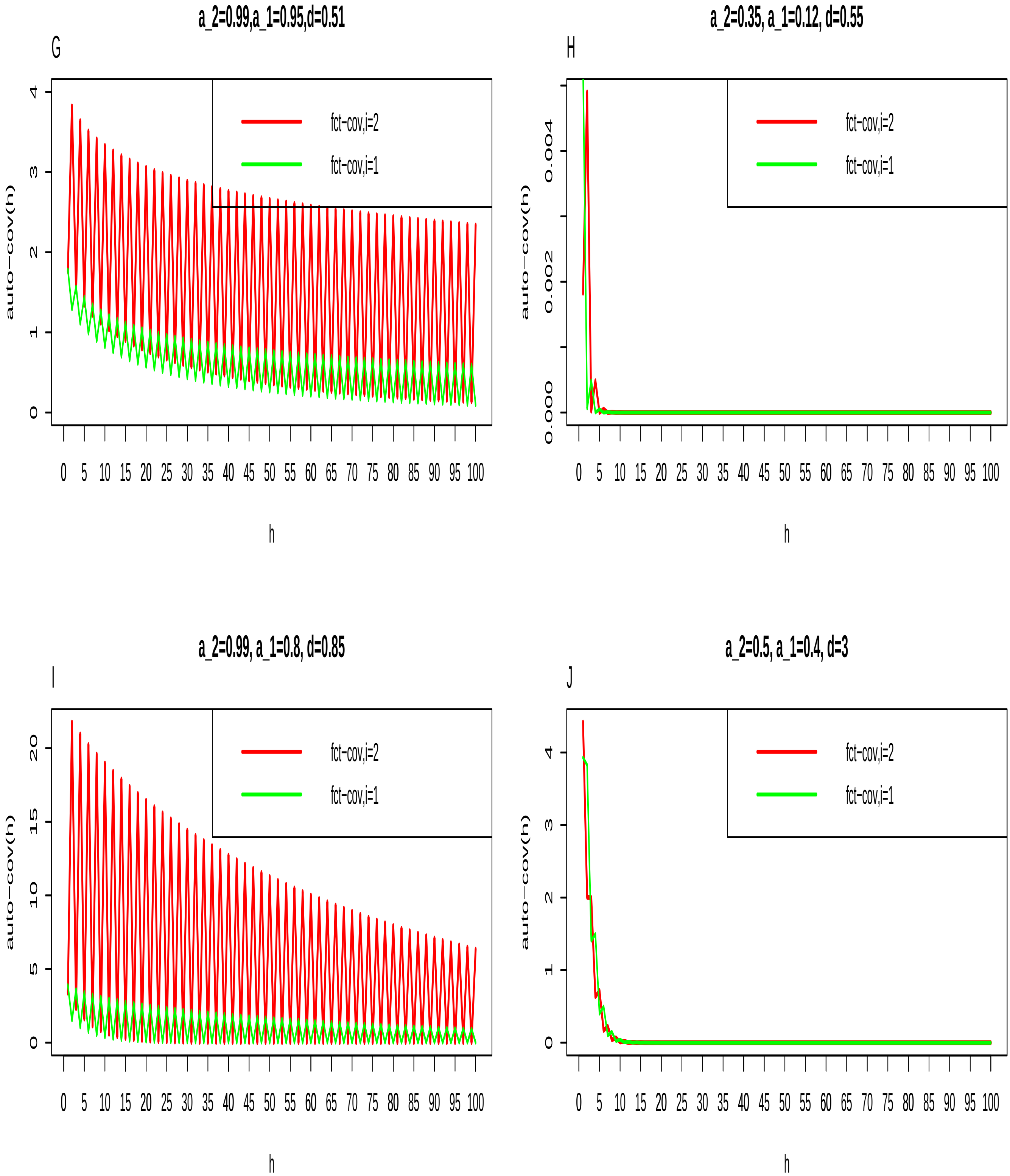

Figure 2.2: Periodic autocovariance $s = 2$.

We remark periodic patterns in subplots A, B, C and D of figure 2.1 and subplots G and I of figure 2.2 caused by values of a_i ($i = 1$ and $i = 2$) near 1, and for $a_2 > a_1$, we can see that $\gamma_x^2 > \gamma_x^1$ in all subplots.
γ_x^i decays more slowly for a value of a_i closer to 1, a hyperbolic form is shown in subplots A, B and G confirming that the function decays more slowly on the boundary of 1 and d around $\frac{1}{2}$, and it decays more rapidly towards 0 for a value of a_i far from 1 and the other values of d.
In subplots C and D, the value of the parameter d is different from A and B. We can see that d affects the behavior of the function, wherein subplot C we observe that the curve of the function decreases nearly in the same way, but if we observe the autocovariance values on the y axis for $h = 0$, it is considerably closer to 0 than in subplot A, and it decays a little faster towards 0, but the periodicity is maintained. As can be seen in the formulas that have been simulated, the autocovariance depends jointly on a_i ($i = 1; 2$) and on the product $a_2 a_1$, and it's evident that the decrease is not always hyperbolic. This observation yields some further results, and to illustrate them we have taken two more examples in subplots E and F, where the value of a_i is smaller than in the previous subplots.
These two subplots confirm the earlier findings. As the periodic coefficient approaches 0 (so that the product between them approaches 0), all the curves decay exponentially towards 0 and we note an absence of any periodicity on the graphs.

Note that a hyperbolic decay of an autocovariance function means that the process models phenomena with a long-memory, whereas an exponential decay indicates that the process models phenomena with a short memory.

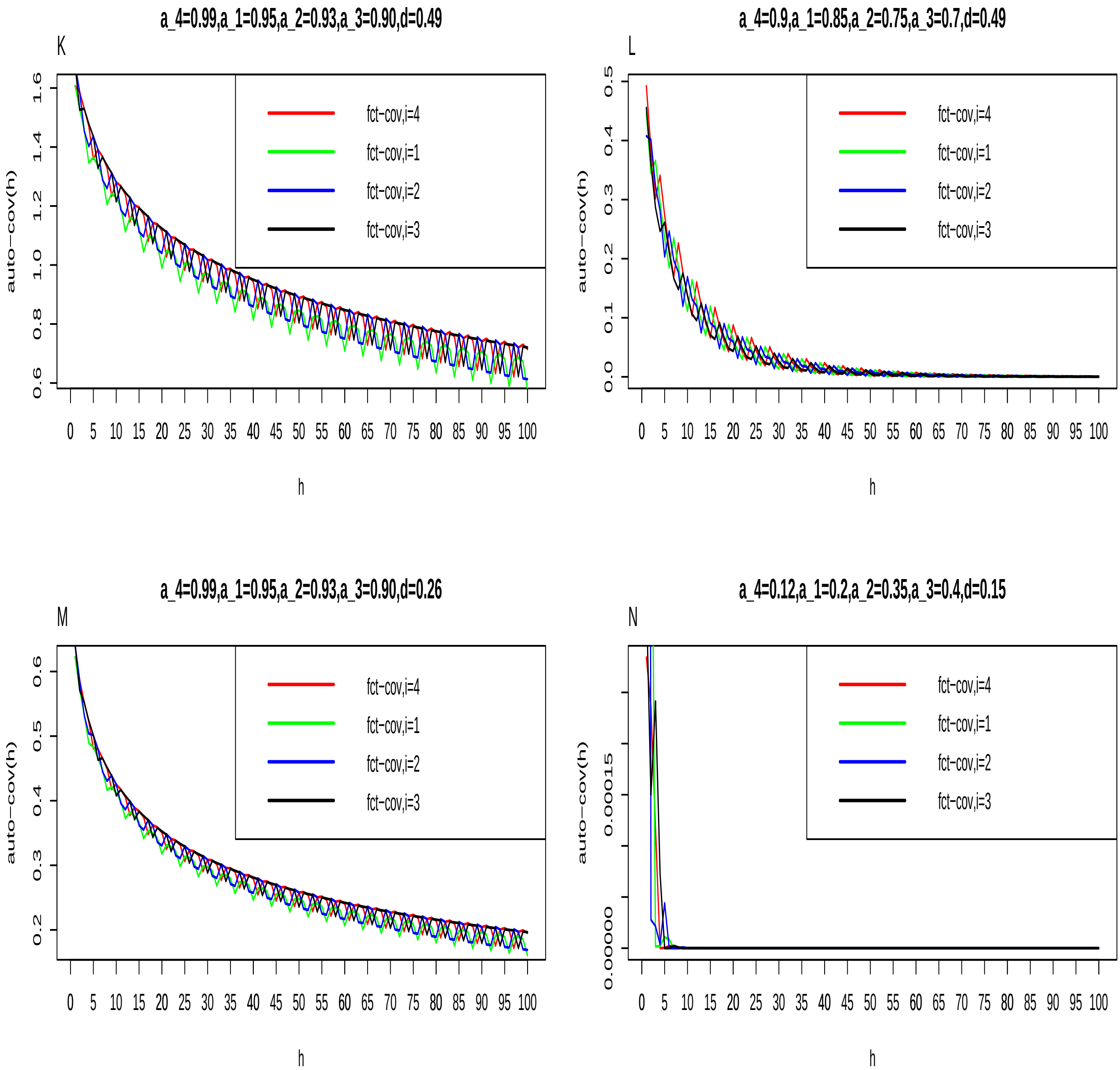

Figure 2.3: Periodic autocovariance $s = 4$.

From the figure 2.3 for a period $s = 4$ we derive similar conclusions to those previously derived for the figure 2.1 and 2.2.

To demonstrate that the results hold for any s, not only for the even cases, we also consider the case $s = 3$; the results are presented below.

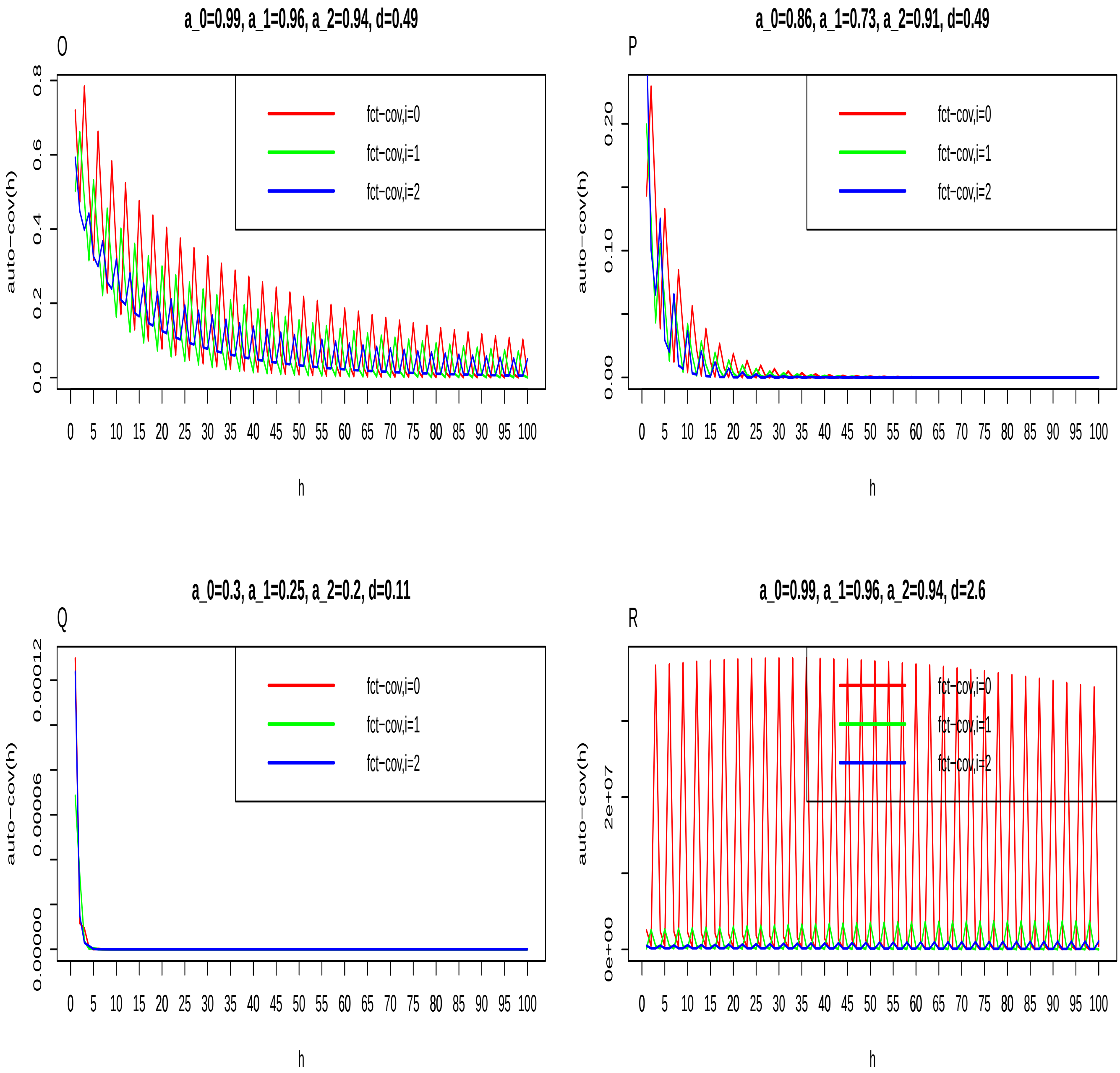

Figure 2.4: Periodic autocovariance $s = 3$.

From Figure 2.4, we deduce that the remarks are similar to the previous cases; thus, the conclusions hold for all periods.

2.5 Conclusion

In this chapter we introduced a new class of fractional autoregressive models characterized by the periodic form of the autocovariance and autocorrelation functions; in other words, we introduced a periodic component to the stochastic equation that defines the fractional autoregressive models, which allows us to model phenomena that can exhibit either short or long-memory, depending on the value of the periodic coefficient. We have shown that the process is non-stationary; rather, it belongs to the known family of periodically stationary processes, and we have defined the forms of the periodic autocorrelation and autocovariance functions together with their asymptotic forms and behavior.
Finally, we examined our main results through a simulation study.

CHAPTER 3

ESTIMATION FOR THE STATIONARY MULTIVARIATE FRACTIONAL AUTOREGRESSIVE PROCESS

3.1 Introduction

The purpose of this chapter is to suggest estimation techniques suitable for non-stationary processes since, on one side, classical methods may not be directly or simply applicable, and on the other side, the estimates asymptotic theory can be challenging to investigate. We apply a method that stands on optimizing the distance between two probability densities; the technique is recognized as a minimum Hellinger distance estimate. We construct the estimation of the parameters for the multivariate fractional autoregressive process, and then under some regularity assumptions, we study its key properties. After that, comparing the estimate with a classical one, we propose another estimate proven to be an analog of the maximum likelihood estimate. The study is well-examined, and both estimates are tested in a simulation approach.

3.2 Stationary multivariate fractional autoregressive model

Let $\{X_t, t \in \mathbb{Z}\}$ be the process satisfying (2.6).
For $i = \{1, ..., s\}$, we put $\theta^{\intercal} = (a_1, ..., a_s, d) \in \Theta \subset \mathbb{R}^{s+1}$, and $\Theta = \{\theta \in \mathbb{R}^{s+1} / \max\limits_{1 \leq i \leq s} |a_i| < 1, d \in \mathbb{R}$ and $a_i.d \neq 0\}$.

We assume that we have a sample $X_t = (X_1, ..., X_n)$ of length n, realizations of the process. Additionally, we assume that n, the sample size, is consistent with the time period s, i.e., $n = sN$; thus, $M = \{0, ..., N-1\}$ such that $t = i + sM$.
$\{X_t, t \in \mathbb{Z}\}$ is a non-stationary but periodically correlated process; however, stationarity is a crucial condition for the sequel. To achieve the stationarity of a periodic process, one solution is to form the multivariate stationary process associated with the initial non-stationary process.
Using the $AR(\infty)$ representation given in (2.5) and for $i = \{1, ..., s\}$, we define the following

$$\eta_M = \begin{pmatrix} \varepsilon_{1+sM} \\ \varepsilon_{2+sM} \\ . \\ . \\ \varepsilon_{s+sM} \end{pmatrix} ; \quad Y_M = \begin{pmatrix} X_{1+sM} \\ X_{2+sM} \\ . \\ . \\ X_{s+sM} \end{pmatrix} \text{ and } \quad \mathit{\Pi}_j = \begin{pmatrix} \Pi_j(1) & 0 & \cdots & 0 \\ 0 & \Pi_j(2) & 0 & \vdots \\ \vdots & 0 & \ddots & 0 \\ 0 & \vdots & 0 & \Pi_j(s) \end{pmatrix},$$

$$\eta_M = \sum_{j=0}^{\infty} \mathit{\Pi}_j Y_{M-j}. \tag{3.1}$$

$\{\eta_M\}$ is the multivariate form of the process defined in (2.5).

3.3 Parameters estimations

In what follows, we describe in detail the estimation techniques we have chosen. We start with the MHDE and then briefly discuss the CSS estimate.

3.3.1 The minimum Hellinger distance estimate

In this subsection, we will define an estimate of θ and let it be denoted by $\hat{\theta}_{M,n}$. This estimator minimizes the Hellinger distance (H^2) between the theoretical probability density of η_M and a random probability density of the estimated η_M (referred to as $\hat{\eta}_M$) denoted by $f_\eta(.)$ and $f_n(.)$ respectively; $\hat{\theta}_{M,n}$ is hence named the **M**inimum **H**ellinger **D**istance (MHD) estimate, and it is defined in the following form

$$\hat{\theta}_{M,n} = arg \min_{\theta \in \Theta} H^2(f_n; f_\eta), \tag{3.2}$$

where $H^2(f_n; f_\eta)$ is the L_2 norm of the deviation between the square-roots of f_n and f_η written as

$$H^2(f_n; f_\eta) = \left(\int_{\mathbb{R}^s} |f_n^{\frac{1}{2}}(x) - f_\eta^{\frac{1}{2}}(x)|^2 dx \right)^{\frac{1}{2}}, \tag{3.3}$$

such that:
$f_\eta(.)$ is the theoretical probability density of η_M, with $f_\eta : \mathbb{R}^s \longrightarrow \mathbb{R}^+$.
$f_n(.)$ is the random function of $\hat{\eta}_M$ given by

$$f_n(x) = \frac{1}{N h_N^s} \sum_{M=0}^{N-1} K(\frac{x - \hat{\eta}_M}{h_N}), \quad x \in \mathbb{R}^s. \tag{3.4}$$

Let $\tilde{f}_n(.)$ be a non-parametric kernel density estimator of f_η defined by

$$\tilde{f}_n(x) = \frac{1}{N h_N^s} \sum_{M=0}^{N-1} K(\frac{x - \eta_M}{h_N}), \quad x \in \mathbb{R}^s, \tag{3.5}$$

where $K : \mathbb{R}^s \longrightarrow \mathbb{R}^+$ is a kernel density function, (h_N) is a sequence of bandwidths, $\hat{\eta}_M$ is defined such that $\hat{\eta}_M = (\hat{\varepsilon}_{1+sM}, ..., \hat{\varepsilon}_{s+sM})^\intercal$ and $\forall i = \{1, ..., s\}$, $\hat{\varepsilon}_{i+sM} = \sum_{j=0}^{n} \Pi_j(i) X_{i+sM-j}$,
and $f_n : \mathbb{R}^s \longrightarrow \mathbb{R}^+$.
According to (3.1), we write the following

$$\hat{\eta}_M = \sum_{j=0}^{n} \mathit{\Pi}_j Y_{M-j}. \tag{3.6}$$

3.3.1.1 Regularity assumptions

To study the main characteristics of $\hat{\theta}_{M,n}$, we need to set some assumptions and conditions on the functions and the components we defined in the previous subsection.

Assumption A.

- $E(||\eta_M||^l) < \infty$ for $l \in \mathbb{N}^*$.
- For all $(u, v) \in \mathbb{R}^{2s}$, we have:
  - ⋆ $\int_{\mathbb{R}^s} K^2(u)du < \infty$; $\int_{\mathbb{R}^s} u_i K(u)du = 0$ for $i = \{1, ..., s\}$.
  - ⋆ $\int_{\mathbb{R}^s} u_i u_j K(u)du = 0$; $\int_{\mathbb{R}^s} u_i^2 K(u)du < \infty$ for $i = \{1, ..., s\}$, $j = \{1, ..., s\}$ and $i \neq j$.
  - ⋆ There exists a positive constant c such that $\sup_{u \in \mathbb{R}^s} |K(u+v) - K(u)| \leq c||v||$. Here $||.||$ denotes any norm on $\mathbb{R}^k$ ($k = s$ or $s + 1$).

Assumption B.

1. η_M admits an absolutely continuous density with respect to the Lebesgue measure on $\mathbb{R}^s$ positive in a neighborhood of the origin.
   For any $\theta \in \Theta$ and each $x \in \mathbb{R}^s$, the functions $x \longmapsto f_\eta(x)$ and $x \longmapsto f_\eta^{\frac{1}{2}}(x)$ are continuously differentiable.
2. For each $x \in \mathbb{R}^s$, the functions $\theta \longmapsto \frac{\partial}{\partial \theta_j} f_\eta^{\frac{1}{2}}(x)$ for $j = \{1, ..., s\}$ and $\theta \longmapsto \frac{\partial^2}{\partial \theta_j \theta_k} f_\eta^{\frac{1}{2}}(x)$ for $1 \leq j; k \leq s$ are bounded, continuous and in $L^2(\mathbb{R}^s)$.

Assumption C. (h_N) the sequence of bandwidths satisfies:

- $h_N = N^\alpha \ell(N), \, where -1 < \alpha < 0$ and $\ell(.)$ is a slowly varing function.
- As N tends to infinity h_N tends to 0, Nh_N tends to infinity and $\lim_{N \to \infty} \frac{\ell(bN)}{\ell(N)} = 1$; $b > 0$.

For each $\theta \in \Theta$, $\sup_{x \in \mathbb{R}^s} |\frac{\partial^j f_\eta}{\partial x_k^j}(x)| < \infty$ with $j \in \mathbb{N}$ and $k = \{1, ..., s\}$.

Assumption D. For $\theta, \theta' \in \Theta$, $\theta \neq \theta'$ implies that $\{x \in \mathbb{R}^s / f_\eta(x) \neq f_{\eta'}(x)\}$ is a set of positive Lebesgue measure.

Assumption E. There exists a finite constant value m such that $\sup_{x \in \mathbb{R}^s} f_n(x) \leq m$.

Based on Mbeke and Hili [75], we define the following:

$$g_\eta(x) = f_\eta^{\frac{1}{2}}(x) \; ; \; g'_\eta(x) = \frac{\partial g_\eta(x)}{\partial \theta} \; ; \; g''_\eta(x) = \frac{\partial^2 g_\eta(x)}{\partial \theta \partial \theta^\intercal} \; ;$$
$$S_\theta(x) = \left[\int_{\mathbb{R}^s} g'_\eta(x)(g'_\eta(x))^\intercal dx \right]^{-1} g'_\eta(x).$$

Condition I. Both g'_η and g''_η have components in L_2 and their norms are continuous functions at θ.

Condition II. $\int_{\mathbb{R}^s} g''_\eta(x) g_\eta(x) dx$ is non-singular $(s+1 \times s+1)$-dimensional matrix.

Remarks on the assumptions

1) Assumptions **A** and **B** are essential technical conditions for establishing the asymptotic properties of the estimator. These assumptions concern the regularity conditions of the kernel, the white noise density function and its l-order moment.

2) The less technical assumptions **C** and **D** concern the smoothing parameter, which is the degree of smoothing determined by the extent of the kernel on either side of the observation. They also determine the conditions under which the parameter can be identified.

3) E is an assumption used in the application of the Prakasa Rao (1983) inequality [86] to demonstrate convergence of the random part of the estimator. This hypothesis assumes the existence of a finite quantity, and f_n denotes a density function defined in the Prakasa Rao inequality.

3.3.1.2 Main results

Lemma 3.3.1 *If the assumptions A and B.1 are fulfilled, then f_n almost surely converges to f_η as $N \longrightarrow \infty$.*

Proof 9 *We have*

$$f_n(x) - f_\eta(x) = f_n(x) - \tilde{f}_n(x) + \tilde{f}_n(x) - \mathbb{E}(\tilde{f}_n(x)) + \mathbb{E}(\tilde{f}_n(x)) - f_\eta(x),$$

by applying the triangular inequality to the previous form, we get

$$\sup_{x\in\mathbb{R}^s} |f_n(x) - f_\eta(x)| \le \sup_{x\in\mathbb{R}^s} |f_n(x) - \tilde{f}_n(x)| + \sup_{x\in\mathbb{R}^s} |\tilde{f}_n(x) - \mathbb{E}(\tilde{f}_n(x))| + \sup_{x\in\mathbb{R}^s} |\mathbb{E}(\tilde{f}_n(x)) - f_\eta(x)|.$$

Below, we shall prove the almost sure convergence of each of the terms on the right side of the above inequality separately.

i) In the first term, we substitute the values defined in section 3.3.1 for each of $f_n(x)$ *and* $\tilde{f}_n(x)$*, and we obtain*

$$\sup_{x\in\mathbb{R}^s}|f_n(x)-\tilde{f}_n(x)| = \frac{1}{Nh_N^s}\sup|\sum_{M=0}^{N-1}K(\frac{x-\hat{\eta}_M}{h_N})-\sum_{M=0}^{N-1}K(\frac{x-\eta_M}{h_N})|,$$

by assumption A, we write

$$\sup_{x\in\mathbb{R}^s}|f_n(x)-\tilde{f}_n(x)| \leq c\frac{1}{Nh_N^{s+1}}\sum_{M=0}^{N-1}||e_M||,$$

where e_M *represents* $\eta_M-\hat{\eta}_M$*, by (3.1) and (3.6), we find*

$$e_M = \sum_{j=n+1}^{\infty}\mathit{\Pi}_j Y_{M-j}, \tag{3.7}$$

according to Granger and Andersen [36], we consider the following expression

$$\begin{aligned}\mathbb{E}(\frac{1}{Nh_N^{s+1}}\sum_{M=0}^{N-1}||e_M||)^2 &= \mathbb{E}(\frac{1}{N^2h_N^{2(s+1)}}\sum_{M=0}^{N-1}||e_M||^2)+2\mathbb{E}[\frac{1}{N^2h_N^{2(s+1)}}\sum_{\substack{M,w=0\\M\neq w}}^{N-1}(||e_M||)(||e_w||)]\\ &\leq \frac{1}{N^2h_N^{2(s+1)}}(2\sum_{M=0}^{N-1}\mathbb{E}||e_M||^2+\sum_{t=0}^{N-1}\mathbb{E}||e_w||^2),\end{aligned}$$

the coefficients $\Pi_j(i)$ *for* $i=\{1,...,s\}$ *are square summable (Odaki [80]), and as per Mbeke and Hili [75],* $\sup_{x\in\mathbb{R}^s}|f_n(x)-\tilde{f}_n(x)|$ *almost surely converges to* 0.

ii) For the second term, we consider the Prakasa Rao's inequality [86] amended by Mbeke and Hili [75] in accordance with the form of the multivariate process.

We set $\delta_M(x)=\frac{1}{h_N^s}K(\frac{x-\eta_M}{h_N})$*, and the sequence* $(s_n)_{n\in\mathbb{N}^*}$ *such that* $s_N=Nh_N$*, there is a constant* $c_0>0$ *such that*

$$\sup_{x\in\mathbb{R}^s}\delta_M(x)\leq c_0 s_N,$$

we write the Prakasa Rao's inequality as follows

$$\mathbb{P}\left[|\tilde{f}_n(x)-\mathbb{E}(\tilde{f}_n(x))|>\varepsilon\sqrt{\frac{s_N log(N)}{N}}\right]\leq 2exp\left(-\frac{s_N log(N)\varepsilon^2}{8c_0 m}\right),$$

under assumption C and the selection of $h_N=N^\alpha log(N)$*, the previous inequality becomes as follows*

$$\mathbb{P}\left[|\tilde{f}_n(x)-\mathbb{E}(\tilde{f}_n(x))|>\varepsilon N^{\frac{\alpha}{2}}log(N)\right]\leq 2exp\left(-\frac{N^{\alpha+1}log^2(N)\varepsilon^2}{8c_0 m}\right),$$

$$\mathbb{P}\left[N^{\frac{1}{4}}\sup_{x\in\mathbb{R}^s}|\tilde{f}_n(x)-\mathbb{E}(\tilde{f}_n(x))|>\varepsilon N^{\frac{2\alpha+1}{4}}log(N)\right]\leq 2exp\left(-\frac{N^{\frac{4\alpha+5}{4}}log^2(N)\varepsilon^2}{8c_0m}\right),$$

since $-1<\alpha<0$, *by assumption E and Mbeke and Hili [75], we find that* $-1<\dfrac{n^{\frac{4\alpha+5}{4}}log^2(n)\varepsilon^2}{8c_0m}<\infty.$

There exists a sequence $w_N=\beta log(N)$, *where* $\beta\geq 2$, *that for some rank it verifies*

$$2exp\left(-\frac{N^{\frac{4\alpha+5}{4}}log^2(N)\varepsilon^2}{8c_0m}\right)<\frac{2}{N^\beta},$$

hence, we can write

$$\mathbb{P}\left[N^{\frac{1}{4}}\sup_{x\in\mathbb{R}^s}|\tilde{f}_n(x)-\mathbb{E}(\tilde{f}_n(x))|>\varepsilon N^{\frac{2\alpha+1}{4}}log(N)\right]\leq\frac{2}{N^\beta},$$

$$\sum_{N\geq 1}\mathbb{P}\left[N^{\frac{1}{4}}\sup_{x\in\mathbb{R}^s}|\tilde{f}_n(x)-\mathbb{E}(\tilde{f}_n(x))|>\varepsilon N^{\frac{2\alpha+1}{4}}log(N)\right]\leq\sum_{N\geq 1}\frac{2}{N^\beta},$$

since $\sum_{N\geq 1}\dfrac{2}{N^\beta}$ *converges, we get*

$$\sum_{N\geq 1}\mathbb{P}\left[N^{\frac{1}{4}}\sup_{x\in\mathbb{R}^s}|\tilde{f}_n(x)-\mathbb{E}(\tilde{f}_n(x))|>\varepsilon N^{\frac{2\alpha+1}{4}}log(n)\right]<\infty.$$

And by Borel-Cantelli's lemma it result that $\sup_{x\in\mathbb{R}^s}|\tilde{f}_n(x)-\mathbb{E}(\tilde{f}_n(x))|$ *converges to zero almost surely when* n *tends to infinity.*

iii) From (3.5) we write

$$\begin{aligned}
\mathbb{E}(\tilde{f}_n(x)) &= \frac{1}{Nh_N^s}\mathbb{E}[\sum_{M=0}^{N-1}K(\frac{x-\eta_M}{h_N})]\\
&= \frac{1}{h_N^s}\mathbb{E}[K(\frac{x-\eta_1}{h_N})]\\
&= \frac{1}{h_N^s}\int_{\mathbb{R}^s}K(\frac{x-z}{h_N})f_\eta(z)dx\\
&= \int_{\mathbb{R}^s}K(u)f_\eta(x-uh_N)du.
\end{aligned}$$

We deduce the following using Taylor's expansion to the second order near x *on* $f_\eta(x-uh_N)$

$$\mathbb{E}(\tilde{f}_n(x))-f_\eta(x)=\int_{\mathbb{R}^s}K(u)[f_\eta(x-uh_N)-f_\eta(x)]du,$$

$$\sup_{u\in\mathbb{R}^s} |\mathbb{E}(\tilde{f}_n(x)) - f_\eta(x)| \leq \frac{h_N^2}{2} \sum_{k=1}^{s} \sup_{x\in\mathbb{R}^s} \left|\frac{\partial^2 f_\eta}{\partial^2 x_k^2}(x)\right| \int_{\mathbb{R}^s} K(u)[u_k^2 + o(h_N^2)]du.$$

By assumption A, we find that $\int_{\mathbb{R}^s} K(u)[u_k^2+o(h_N^2)]du < \infty$ *and by assumption C we have* $\sup_{x\in\mathbb{R}^s}\left|\frac{\partial^2 f_\eta}{\partial^2 x_k^2}(x)\right| < \infty$*, hence* $\sup_{x\in\mathbb{R}^s} |\mathbb{E}(\tilde{f}_n(x)) - f_\eta(x)| \longrightarrow 0$ *as* n *tends to infinity. After showing the convergence of the three terms, it becomes clear that* f_n *almost surely converges to* f_η *as* $N \longrightarrow \infty$.

Lemma 3.3.2 *Let* F *be the set containing all densities metrized by the* L_1 *distance.* $T : F \longrightarrow \Theta$ *is the functional, and let us determine a density* $g \in F$ *and* $A(g)$ *a set such that*

$$A(g) = \{\theta \in \Theta : H^2(g; f_\eta)|_{\theta=\theta_0} = \min_{\theta\in\Theta} H^2(g; f_\eta)\}.$$

Since $A(g)$ *may include multiple values, we assume that it stands for any arbitrary but single element, and we consider that element to be* $T(g)$.

Proof 10 *We refer to the proofs of Beran [8] and Hili [48].*

Lemma 3.3.3 *Under the assumptions B1 and C and the conditions I and II, and supposing that* θ *is within* Θ*; thus, for all sequences* $\{f_n\}$ *converge to* f_η *in the Hellinger metric, we get that*

$$T(f_n(x)) = \theta + \int_{\mathbb{R}^s} S_\theta(x)\left(f_n^{\frac{1}{2}}(x) - f_\eta^{\frac{1}{2}}(x)\right)dx + G_N \int_{\mathbb{R}^s} g'_\eta(x)\left(f_n^{\frac{1}{2}}(x) - f_\eta^{\frac{1}{2}}(x)\right)dx,$$

where G_N *is a non-singular* $(s + 1 \times s + 1)$*-matrix whose components tend to zero after multiplication by* $\sqrt{N}$*, i.e.,* $(\sqrt{N}G_N) \underset{N\longrightarrow\infty}{\longrightarrow} 0$.

Proof 11 *This result is clearly well demonstrated by Beran [8].*

Lemma 3.3.4 *Provided that assumptions A, B1 and C are true, then*

$$\sqrt{Nh_N}(f_n(x) - f_\eta(x)) \underset{N\rightarrow\infty}{\longrightarrow} \mathbb{N}\left(0, f_\eta(x)\int_{\mathbb{R}^s} K^2(u)du\right).$$

Proof 12 *This proof corresponds with the one provided by Wu and Mielniczuk [109] for this result.*

3.3.1.3 Estimator properties

Theorem 3.1 (Consistency). *Considering that all the prior assumptions are verified, $\hat{\theta}_{M,n}$ almost surely converges to θ when $N \longrightarrow \infty$.*

Proof 13 *By lemma 3.3.1, since $|f_n(x) - f_\eta(x)| \underset{N\longrightarrow\infty}{\longrightarrow} 0$ almost surely, then*

$$\mathbb{P}\{\lim_{N\longrightarrow\infty} f_n^{\frac{1}{2}}(x) = f_\eta^{\frac{1}{2}}(x) \quad \forall x\} = 1,$$

therefore

$$H^2(f_n; f_\eta) \underset{N\longrightarrow\infty}{\longrightarrow} 0 \quad \text{almost surely.}$$

And from lemma 3.3.2 the functional $T(f_\eta) = \theta$ unique on Θ, so T is continuous at f_η in the Hellinger topology, thus

$$\hat{\theta}_{M,n} = T(f_n(x)) \underset{N\longrightarrow\infty}{\longrightarrow} T(f_\eta(x)) = \theta.$$

i.e., $\hat{\theta}_{M,n} \underset{N\longrightarrow\infty}{\longrightarrow} \theta$ almost surely.

Theorem 3.2 (Asymptotic distribution). *Assuming that assumptions A to E hold true and if both conditions I and II are verified, then*

$$\sqrt{N}(\hat{\theta}_{M,n} - \theta) \underset{N\longrightarrow\infty}{\longrightarrow} \mathbb{N}(0, \Sigma^2),$$

where Σ^2 the variance-covariance matrix is given by

$$\frac{1}{4}\left[\int_{\mathbb{R}^s} g'_\eta(x)(g'_\eta(x))^\intercal dx\right]^{-1} \int_{\mathbb{R}^s} K^2(u)du.$$

Proof 14 *As $T(f_n(x)) = \hat{\theta}_{M,n}$ and according to lemma 3.3.3 we have*

$$\sqrt{N}(\hat{\theta}_{M,n}-\theta) = \sqrt{N}\int_{\mathbb{R}^s} S_\theta(x)(f_n^{\frac{1}{2}}(x)-f_\eta^{\frac{1}{2}}(x))dx+\sqrt{N}G_N\int_{\mathbb{R}^s} g'_\eta(x)(f_n^{\frac{1}{2}}(x)-f_\eta^{\frac{1}{2}}(x))dx,$$

and $\sqrt{N}G_N \underset{N\longrightarrow\infty}{\longrightarrow} 0$ *thence*

$$\sqrt{N}(\hat{\theta}_{M,n}-\theta) = \sqrt{N}\int_{\mathbb{R}^s} S_\theta(x)(f_n^{\frac{1}{2}}(x) - f_\eta^{\frac{1}{2}}(x))dx + o_p(1),$$

$S_\theta(x) \in L_2$, *furthermore,* S_θ *and* $f_\eta^{\frac{1}{2}}$ *are orthogonal in* L_2.

Under assumption B.1, since $f_\eta^{\frac{1}{2}}$ *is positive and by the following algebraic equality*

$$f_n^{\frac{1}{2}}(x) - f_\eta^{\frac{1}{2}}(x) = \frac{f_n(x)-f_\eta(x)}{2f_\eta^{\frac{1}{2}}(x)} - \frac{(f_n(x)-f_\eta(x))^2}{2f_\eta^{\frac{1}{2}}(x)(f_n^{\frac{1}{2}}(x)-f_\eta^{\frac{1}{2}}(x))^2},$$

we can write

$$\sqrt{n}(\hat{\theta}_{M,n}-\theta) = \sqrt{N}\int_{\mathbb{R}^s} S_\theta(x)\frac{f_n(x)-f_\eta(x)}{2f_\eta^{\frac{1}{2}}(x)}dx + r_n, \tag{3.8}$$

such as

$$r_n = -\sqrt{N}\int_{\mathbb{R}^s} S_\theta(x)\frac{(f_n(x)-f_\eta(x))^2}{2f_\eta^{\frac{1}{2}}(x)(f_n^{\frac{1}{2}}(x)-f_\eta^{\frac{1}{2}}(x))^2}dx,$$

$$||r_n|| \leq \sqrt{N}\int_{\mathbb{R}^s} ||S_\theta(x)||\frac{(f_n(x)-f_\eta(x))^2}{2f_\eta^{\frac{1}{2}}(x)(f_n^{\frac{1}{2}}(x)-f_\eta^{\frac{1}{2}}(x))^2}dx.$$

Let's set $\xi = \inf_{x\in\mathbb{R}^s} f_\eta(x)$, *by applying the following inequality*

$$2f_\eta^{\frac{1}{2}}(x)(f_n^{\frac{1}{2}}(x)-f_\eta^{\frac{1}{2}}(x))^2 = 2f_\eta^{\frac{3}{2}}(x)+\gamma > 2f_\eta^{\frac{3}{2}}(x), \quad when \quad \gamma > 0,$$

we find

$$||r_n|| \leq \frac{1}{2\xi^{\frac{3}{2}}}\int_{\mathbb{R}^s} ||S_\theta(x)||\sqrt{N}(f_n(x)-f_\eta(x))^2dx.$$

$S_\theta(x)$ *is continuous and bounded (according to conditions I and II), by lemma 3.3.1 and Vitali's convergence theorem,* $||r_n||$ *converges in probability to* 0 *as* $N \longrightarrow \infty$.

Now (3.8) can be expressed as

$$\sqrt{N}(\hat{\theta}_{M,n}-\theta) = \sqrt{N}\int_{\mathbb{R}^s} S_\theta(x)\frac{f_n(x)-f_\eta(x)}{2f_\eta^{\frac{1}{2}}(x)}dx + o_p(1),$$

therefrom and by lemma 3.3.4 $\sqrt{N}(\hat{\theta}_{M,n}-\theta) \longrightarrow \mathbb{N}(0;\Sigma^2)$.

And

$$\begin{aligned}
\Sigma^2 &= \int_{\mathbb{R}^s}\left(\frac{S_\theta(x)}{2f_\eta^{\frac{1}{2}}(x)}\right)\left(\frac{S_\theta(x)}{2f_\eta^{\frac{1}{2}}(x)}\right)^\intercal dx \int_{\mathbb{R}^s} K(u)^2du f_\eta(x)dx \\
&= \int_{\mathbb{R}^s}\frac{1}{4}(S_\theta(x))(S_\theta(x))^\intercal dx \int_{\mathbb{R}^s} K(u)^2du \\
&= \frac{1}{4}\left[\int_{\mathbb{R}^s} g'_\eta(x)(g'_\eta(x))^\intercal dx\right]^{-1}\int_{\mathbb{R}^s} K^2(u)du.
\end{aligned}$$

3.3.2 The conditional sum of squares estimate

Considering the multivariate stationary process $\{Y_M, M \in \mathbb{Z}\}$ given in (2.7) related to $\{X_t, t \in \mathbb{Z}\}$ provided in (2.6) and assuming that the innovation process $\{\varepsilon_t, t \in \mathbb{Z}\}$ are independent and identically distributed with the Gaussian distribution, it might be a possibility to use maximum likelihood estimation (MLE) to obtain the estimated values of the parameters of $\{Y_M, M \in \mathbb{Z}\}$. But the form of the variance-covariance matrix of $\{Y_M, M \in \mathbb{Z}\}$ makes this method difficult. Box and Jenkins [10] demonstrated that the maximum likelihood estimates are the equivalent of the least squares estimates derived by maximizing the conditional sum of function; this approach has been widely applied in estimation in various studies; for long memory models, we can refer to Li and McLeod [73] and Chung [18]. In this paragraph, we propose to estimate the parameters of our process using the same procedure.

We refer to the **conditional sum of squares (CSS) estimates** by $\hat{\theta}$, which is the value that minimizes the sum of squared innovations.
The expression $\hat{\theta}$ is represented in the form below

$$\hat{\theta} = \arg\min_{\theta \in \Theta} \sum_{M=0}^{N-1} ({\eta_M}^{\intercal})\eta_M, \tag{3.9}$$

and

$$\eta_M = \sum_{j=0}^{\infty} \mathit{\Pi}_j Y_{M-j},$$

where

$$\mathit{\Pi}_j = \begin{pmatrix} \Pi_j(1) & 0 & \cdots & 0 \\ 0 & \Pi_j(2) & 0 & \vdots \\ \vdots & 0 & \ddots & 0 \\ 0 & \vdots & 0 & \Pi_j(s) \end{pmatrix}.$$

Since η_M is not observable, we can estimate it by $\hat{\eta}_M = (\hat{\varepsilon}_{1+sM}, \hat{\varepsilon}_{2+sM}, ..., \hat{\varepsilon}_{s+sM})^{\intercal}$, such that $\hat{\varepsilon}_{i+sM} = \sum_{j=0}^{n} \Pi_j(i) X_{i+sM-j}$ for all $i \in \{1, ..., s\}$.

3.4 Simulation

We dedicate this section to the validation of the estimator's theoretical results. We present some simulated values related to the MHDE $\hat{\theta}_{M,n}$ given by the formula (3.2) and the outcomes of the CSS estimate $\hat{\theta}$ given in the formula (3.9).

We begin with the MHD estimate, we consider the case of a period s equals 2; we assume that the white noise η_M is the density function of the standard normal distribution, the kernel density K is similarly the density function of the standard normal distribution, and we select an (h_N) that satisfies assumption C.

For different sample sizes n and multiple $\theta^{\intercal} = (a_1, a_2, d)$ values, and for $R = 1000$ replications, we calculate:

- R values of $\hat{\theta}_{M,n}$, we denote $(\hat{\theta}^{(r)}_{M,n})^{\intercal} = (\hat{a}^{(r)}_1, \hat{a}^{(r)}_2, \hat{d}^{(r)})$ with $r = 1, ..., R$.
- The mean of the sample of $\hat{\theta}_{M,n}$, denoted $\hat{\hat{\theta}}^{\intercal}_{M,n} = (\hat{\hat{a}}_1, \hat{\hat{a}}_2, \hat{\hat{d}})$.
- The general mean squared errors MSE of $\hat{\theta}_{M,n}$ by the following formula

$$MSE = \frac{1}{R}\sum_{r=1}^{R} ||\hat{\theta}^{(r)}_{M,n} - \theta||^2.$$

- The individual confidence intervals CI of θ components, such that $CI_\alpha(a_1)$ for a_1, $CI_\alpha(a_2)$ for a_2, and $CI_\alpha(d)$ for d. Since the sample size R is large, the CI will be of the following form $CI_\alpha(\theta_j) = \left[\hat{\hat{\theta}}_{M,n,j} - \frac{z_{\frac{\alpha}{2}}}{\sqrt{R}}\sqrt{var(\hat{\theta}_{M,n,j})}, \hat{\hat{\theta}}_{M,n,j} + \frac{z_{\frac{\alpha}{2}}}{\sqrt{R}}\sqrt{var(\hat{\theta}_{M,n,j})}\right]$.

θ_j (respectively $\theta_{M,n,j}$) represents the j^{th} component of θ (respectively $\theta_{M,n}$), $j = 1, \cdots, s+1$.

$z_{\frac{\alpha}{2}}$ is the critical value from the normal distribution for $(1-\alpha)$ confidence level; we tack $\alpha = 0.05$, $z_{0.025} = 1.96$.

$var(\hat{\theta}_{M,n,j})$ is the j^{th} element of the main diagonal of the variance-covariance matrix of $\hat{\theta}_{M,n}$.

Additionally, we provide a simultaneous representation of the CI as a cube that defines the limits of the individual CI.

The results are stated below (the cubes are from left to right for $n = 50, 100, 150$, respectively).

n	50	100	150
$\hat{\hat{\theta}}^{\mathsf{T}}_{M,n}$	(.0915293;.349688;.165463)	(.106936;.35398;.163461)	(.098171;.349247;.143044)
MSE	0.001004423	0.0004860885	0.0002301472
$CI_\alpha(a_1)$, $CI_\alpha(a_2)$, $CI_\alpha(d)$	[0.09953647,0.1344709], [0.33517246,0.3781367], [0.13624799,0.1699082]	[0.09493727,0.1224020], [0.34443183,0.3727093], [0.14539888,0.1662093]	[0.09021311,0.1084286], [0.34290543,0.3616760], [0.12693061,0.1454963]

Table 3.1: **MHDE results for** $\theta^{\mathsf{T}} = (.1; .35; .15)$ **and a normally distributed** K

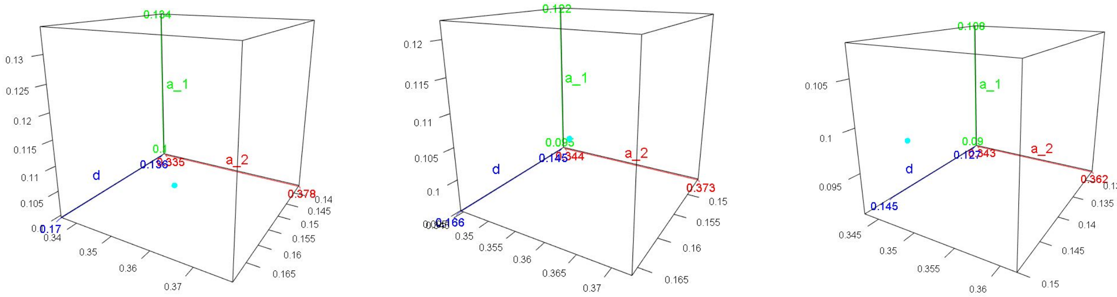

Figure 3.1: Confidence intervals cubes for $\theta^{\mathsf{T}} = (.1; .35; .15)$ and a normally distributed K

n	50	100	150
$\hat{\hat{\theta}}^{\mathsf{T}}_{M,n}$	(.6159309 ;.9119052;.500223)	(.60832;.896313;.49846)	(.602385;.89405;.489826)
MSE	0.0007990064	0.0005892269	0.000179205
$CI_\alpha(a_1)$, $CI_\alpha(a_2)$, $CI_\alpha(d)$	[0.6037686,0.6316076], [0.8997790,0.9356688], [0.4723367,0.5044660]	[0.5958097,0.6244623], [0.8855625,0.9173829], [0.4740730,0.5021804]	[0.5952092,0.6116362], [0.8883418,0.9052387], [0.4761608,0.4919050]

Table 3.2: **MHDE results for** $\theta^{\mathsf{T}} = (.6; .9; .49)$ **and a normally distributed** K

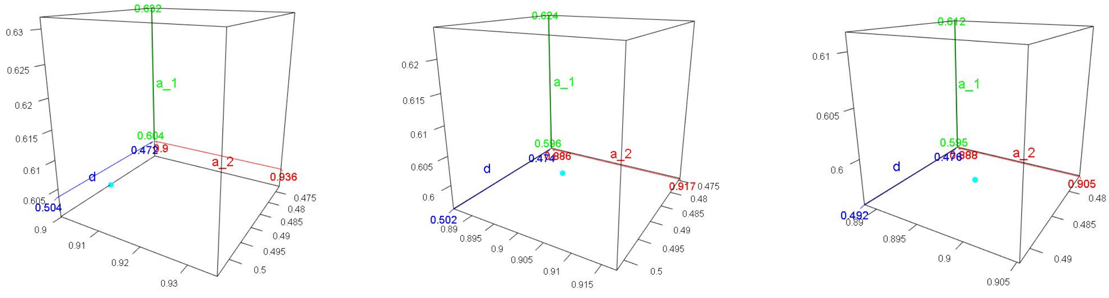

Figure 3.2: Confidence intervals cubes for $\theta^{\intercal} = (.6; .9; .49)$ and a normally distributed K

n	50	100	150
$\hat{\hat{\theta}}^{\intercal}_{M,n}$	(.852387;.73773;3.213104)	(.842364;.74036;3.19282)	(.84017;.74565;3.20007)
MSE	0.0007018098	0.0005319591	0.0002590052
$CI_{\alpha}(a_1)$, $CI_{\alpha}(a_2)$, $CI_{\alpha}(d)$	[0.8413800,0.8665756], [0.7248674,0.7629635], [3.1911542,3.2164435]	[0.8349771,0.8566419], [0.7260081,0.7558865], [3.1923988,3.2162924]	[0.8371869,0.8560224], [0.7370623,0.7510084], [3.1902593,3.2045828]

Table 3.3: **MHDE results for** $\theta^{\intercal} = (.84; .75; 3.2)$ **and a normally distributed** K

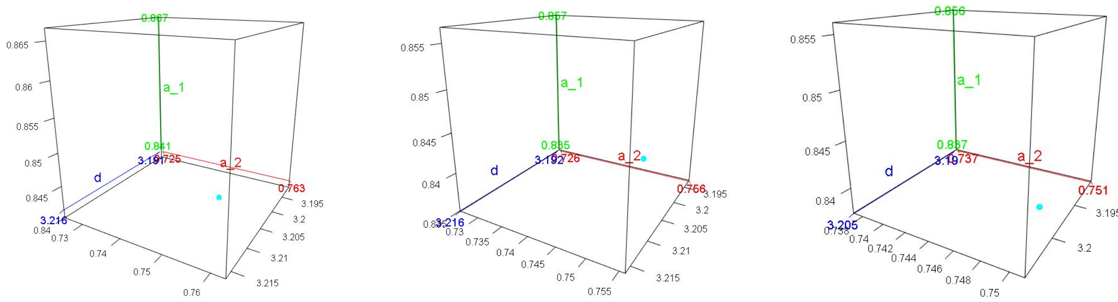

Figure 3.3: Confidence intervals cubes for $\theta^{\intercal} = (.84; .75; 3.2)$ and a normally distributed K

We can clearly see that the mean value of $\hat{\theta}_{M,n}$ is very close to the real parameter in all previous tables (3.1, 3.2, 3.3); this indicates that for a large number of

replications, all estimated values approach θ.
Moreover, we deduce empirically the consistency of the MHD estimate from the mean squared error values.

The MSE decreases significantly as the sample size increases, i.e., for a large n, the MSE is negligible, indicating that $\hat{\theta}_{M,n}$ converges to θ.
We can also note a variation in the length of the confidence intervals; the CI reduces as n increases; graphically, in figures 3.1-3.3, θ (presented as the cyan dot) is inside the cubes regardless of the interval length, which expresses a significant precision for θ.

We evaluate an additional scenario in which we assume that the density function of the gamma distribution is the density function of both η_M and K. This allows us to further analyze the theoretical findings of our study. As in tables 3.1 to 3.3, we consider the same θ values.
Below we present the outcomes.

n	50	100	150
$\hat{\hat{\theta}}^{\intercal}_{M,n}$	(.0914218;.347718;.1443874)	(.0934448;.347829;.146415)	(.102799;.352991;.152659)
MSE	0.0008142538	0.0005513384	0.0001727761
$CI_\alpha(a_1)$, $CI_\alpha(a_2)$, $CI_\alpha(d)$	[0.07633857,0.1108635], [0.33516192,0.3723268], [0.11426422,0.1489700]	[0.08136366,0.1090169], [0.33783885,0.3674083], [0.12000550,0.1504329]	[0.09730022,0.1098878], [0.34748274,0.3637868], [0.13603678,0.1551888]

Table 3.4: **MHDE results for $\theta^{\intercal} = (.1; .35; .15)$ and a gamma K**

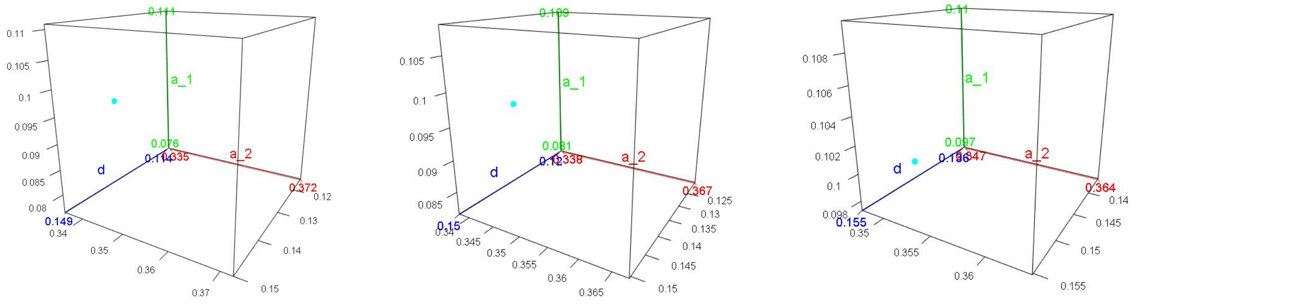

Figure 3.4: Confidence intervals cubes for $\theta^{\intercal} = (.1; .35; .15)$ and a gamma K

n	50	100	150
$\hat{\hat{\theta}}^{\intercal}_{M,n}$	(.5925406;.888854;.490948)	(.600531;.8964;.4921605)	(.600208;.901463;.490411)
MSE	0.0005929565	0.0003468448	6.74331e-05
$CI_{\alpha}(a_1)$, $CI_{\alpha}(a_2)$, $CI_{\alpha}(d)$	[0.5786349,0.6104643], [0.8779601,0.9102040], [0.4703512,0.4940820]	[0.5906979,0.6132061], [0.8896919,0.9096938], [0.4690484,0.4956765]	[0.5965537,0.6049186], [0.8981432,0.9079712], [0.4797295,0.4920365]

Table 3.5: **MHDE results for $\theta^{\intercal} = (.6; .9; .49)$ and a gamma K**

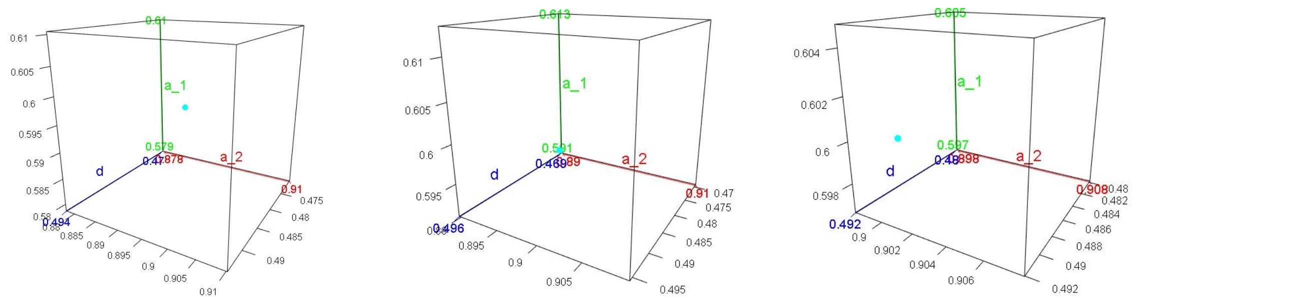

Figure 3.5: Confidence intervals cubes for $\theta^{\intercal} = (.6; .9; .49)$ and a gamma K

n	50	100	150
$\hat{\hat{\theta}}^{\intercal}_{M,n}$	(.837687;.731229;3.21497)	(.841144;.746125;3.19943)	(.835225;.7513308;3.20763)
MSE	0.001115652	0.0002827665	0.0001305248
$CI_\alpha(a_1)$, $CI_\alpha(a_2)$, $CI_\alpha(d)$	[0.8206266,0.8596789], [0.7193254,0.7545566], [3.1798340,3.2203245]	[0.8333723,0.8511633], [0.7393604,0.7593832], [3.1780607,3.2026862]	[0.8318702,0.8409218], [0.7488118,0.7626761], [3.1898919,3.2039336]

Table 3.6: **MHDE results for $\theta^{\intercal} = (.84; .75; 3.2)$ and a gamma K**

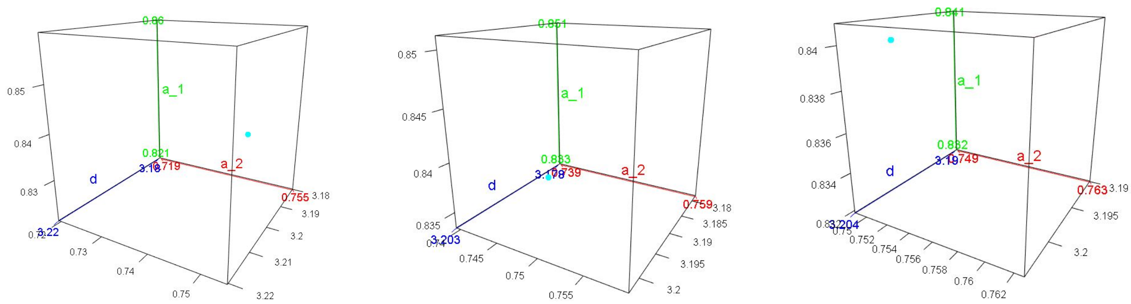

Figure 3.6: Confidence intervals cubes for $\theta^{\intercal} = (.84; .75; 3.2)$ and a gamma K

The results presented in tables 3.4, 3.5 and 3.6 and figures 3.4-3.6 affirm the findings of the first case; however, we note that the $\hat{\hat{\theta}}_{M,n}$ is much closer to the θ and the MSE values decrease faster with the increase of n values compared to the first case.
We can also extract the remark that the results we found hold valid regardless of whether the kernel density was symmetric or asymmetric.

We now examine the validity of the CSS estimate and compare the simulation results with those obtained for the MHD estimate to compare the two estimation techniques.
To do this we generate n realizations of our process $\{Y_M, M \in \mathbb{Z}\}$ for the same values of θ considered in the simulation of the MHD estimate, for the same number of replications R we compute the mean of the R values of $\hat{\theta}^{\intercal} = (\hat{a}_1, \hat{a}_2, \hat{d})$ denoted $\hat{\hat{\theta}}^{\intercal}$. We also calculate the quadratic mean squared errors MSE between θ and $\hat{\hat{\theta}}$.

The results are given in the following tables.

n	50	100	150
$\hat{\hat{\theta}}^{\intercal}$	(.09621508;.3644193;.145004)	(.09985801;.339241;.149241)	(.09929403;.354354;.153548)
MSE	0.001079063	0.0007296985	0.0001980501

Table 3.7: **CSS estimate results for** $\theta^{\intercal} = (.1; .35; .15)$

n	50	100	150
$\hat{\hat{\theta}}^{\intercal}$	(.601191;.909272;.4892333)	(0.591524;0.902044;0.488197)	(.598929;.8971453;.4875943)
MSE	0.0007222061	0.0004311662	0.0001221428

Table 3.8: **CSS estimate results for** $\theta^{\intercal} = (.6; .9; .49)$

n	50	100	150
$\hat{\hat{\theta}}^{\intercal}$	(.835139;.741194;3.206637)	(.8338738;.7567529;3.209463)	(.8355105;.7533341;3.198887)
MSE	0.0008753606	0.0004821385	0.0003217337

Table 3.9: **CSS estimate results for** $\theta^{\intercal} = (.84; .75; 3.2)$

In tables 3.7, 3.8 and 3.9 we can easily observe that the estimated values $\hat{\hat{\theta}}$ are very close to the true parameters θ in all sample sizes, meaning that all $\hat{\theta}$ are close to θ, and that the MSE diminishes noticeably as the sample size gets larger, confirming the validity of the estimation.

Overall, we remark that the MSE of both estimates is almost decreasing similarly.

We note that the CSS estimate can be better or worse than the MHD estimate, indeed the MHD estimate depends on other crucial components, mainly the bandwidth.

3.5 Conclusion

In summary, this chapter aimed to minimize the Hellinger distance between two probability densities to construct an estimate for the parameters of a periodic fractional autoregressive model. We have shown that the estimate fulfills the consistency property and asymptotically follows a normal distribution.
We have also proposed another estimation technique to compare the proposed estimation method with a classical one that is suitable for the estimation of non-stationary models.

Finally, we conducted a simulation study to assess the validity of the estimation results and to compare the two estimations we established in the current study.

CHAPTER 4

LOCAL ASYMPTOTIC PROPERTIES FOR A PERIODIC FRACTIONAL AUTOREGRESSIVE MODEL

4.1 Introduction

In this chapter, we aim to establish local asymptotic normality (LAN), local asymptotic linearity and local asymptotic minimaxity (LAM) for the periodic fractional autoregressive model discussed in the previous chapters. We mainly concentrate on the local asymptotic normality property, which sets the foundation for demonstrating the existence of the local asymptotic minimax estimator and local asymptotic linearity. To complete the chapter and illustrate the LAN results, we present a simulation research.

4.2 Definitions and formulations

4.2.1 Notations and hypotheses

We consider the following series of fractional models

$$(1 - a_i^{(n)} L)^{d^{(n)}} X_{i+sM} = \varepsilon_{i+sM}, \tag{4.1}$$

where $\theta_i^{(n)} = (a_i^{(n)}, d^{(n)})$ defines the local sequence resulting from substituting $a_i^{(n)}$ for a_i and $d^{(n)}$ for d in (2.6) with $a_i^{(n)} = a_i + n^{-\frac{1}{2}}\alpha_i^{(n)}$ and $d^{(n)} = d + n^{-\frac{1}{2}}\delta^{(n)}$, and we define $\tau_i^{(n)} = (\alpha_i^{(n)}, \delta^{(n)})$ to be a sequece of the local perturbation of θ_i such that $Sup||\tau_i^{(n)}|| < \infty$ ($||.||$ denotes the any norm on $\mathbb{R}^2$), $i = 1, \ldots s$. We denote $\theta = (a_1, ..., a_s, d) \in \mathbb{R}^{s+1}$ and $\theta^{(n)} = (a_1^{(n)}, ..., a_s^{(n)}, d^{(n)}) \in \mathbb{R}^{s+1}$.
Let $\mathbb{P}_\theta$ be the probability distribution of the random vector $X = (X_1, ..., X_n)^\intercal$ under θ and $f(.)$ the unspecified probability density of $\{\varepsilon_t\}$.
Let $A_n(X)$ be the $\sigma-$field generated by $(X_t, t \leq n)$, $A_n(\varepsilon)$ the $\sigma-$field generated by $(\varepsilon_t, t \leq n)$ and $A_{n,t}(X)$ the sub $\sigma-$field of $A_n(X)$ generated by the past of the process until the moment t $(X_k, k \leq t, t = 1, ..., n)$.

We suppose that we have a finite length realization $X^{(n)} = (X_1^{(n)}, ..., X_n^{(n)})^\intercal$ of the solution of (2.6) and note by $H_f^{(n)}(\theta)$ the sequence of null hypotheses under which $X^{(n)}$ is generated by (2.6) and $H_f^{(n)}(\theta^{(n)})$ the sequence of alternative hypotheses under which the sequence $X^{(n)}$ is realization of a process satisfying (4.1).
The moving average and autoregressive representations of the process satisfying equation (4.1) are given in equations (4.2) and (4.3), respectively.

$$X_{i+sM}^{(n)} = \sum_{j=0}^{\infty} \psi_j(\theta_i^{(n)}) \varepsilon_{i+sM-j}, \tag{4.2}$$

where

$$\psi_j(\theta_i^{(n)}) = \frac{\Gamma(j + d + n^{-\frac{1}{2}}\delta^{(n)}))}{\Gamma(d + n^{-\frac{1}{2}}\delta^{(n)})\Gamma(j+1)} (a_i + n^{-\frac{1}{2}}\alpha_i^{(n)})^j,$$

and

$$\varepsilon_{i+sM}(\theta_i^{(n)}) = \sum_{j=0}^{\infty} \pi_j(\theta_i^{(n)}) X_{i+sM-j}^{(n)}, \tag{4.3}$$

with

$$\pi_j(\theta_i^{(n)}) = \frac{\Gamma(j - d - n^{-\frac{1}{2}}\delta^{(n)}))}{\Gamma(-d - n^{-\frac{1}{2}}\delta^{(n)})\Gamma(j+1)}(a_i + n^{-\frac{1}{2}}\alpha_i^{(n)})^j.$$

We set the condition that $\sum_{j=1}^{\infty} \psi_j(\theta_i^{(n)})$ and $\sum_{j=1}^{\infty} \pi_j(\theta_i^{(n)})$ are absolutely summable, this condition follows from the one formulated by [94] for a fixed i, $(i = 1, ..., s)$.

4.2.2 Technical regularity assumptions

For the rest of this chapter we need to set some technical regularity assumptions on the density f as follows:

Assumption 1. The process given in (2.6) satisfies a sufficient condition of causality and invertibility that $\forall i = 1, ..., s$, $|a_i| < 1$ and $d \in \mathbb{R}$.

Assumption 2. $\mathbb{E}(\varepsilon_t) = 0$, $\mathbb{E}\left(\varepsilon_t^2\right) = \sigma^2 < \infty$ and $\mathbb{E}\left(\varepsilon_t^4\right) < \infty$.

Assumption 3. f is absolutely continuous over finite intervals (Hàjek and Siddak (1967) [43]), i.e. there exists an integrable function f' on all bounded intervals, such that, for all $-\infty < a < b < \infty$, we have, $f(b) - f(a) = \int_a^b f'(x)dx$.

Assumption 4. Let $\phi_f(.) = \dfrac{-f'(.)}{f(.)}$, and we assume that $0 < \int_a^b |\phi_f(x)|^{2+\delta} dx < \infty$, for some $\delta > 0$ $\left(0 < I(f) = \mathbb{E}\left(\phi_f^2(\varepsilon_t)\right) < \infty\right)$. $I(f)$ represents the Fisher information.

Assumption 5. f is unimodal, i.e., $-log(f(x))$ function is convex on the open intervals $]a, b[$, $-\infty < a < b < \infty$ or $\phi_f(x)$ is increasing on $\mathbb{R}$.

From these assumptions, we deduce:

1) Under assumptions 2 and 3, according to Hàjek and Siddak (1967 [43]) we have,

$$\mathbb{E}(\phi_f(\varepsilon_t)) = \int_R \phi_f(x)f(x)dx = 0,$$

and

$$\mathbb{E}(\varepsilon_t\phi_f(\varepsilon_t)) = \int_R x\phi_f(x)f(x)dx = 1.$$

2) Under assumption 3, it follows that the function $f^{1/2}$ is differentiable in root mean square and we get

$$\underset{t\to 0}{Lim}\frac{1}{t^2}\int_R\left(f^{1/2}(z+t)-f^{1/2}(z)-\frac{t}{2}\frac{f'(z)}{f^{1/2}(z)}\right)^2dz=0.$$

4.2.3 Likelihood ratio sequence

In order to prove the LAN property and therefore pass to the properties of local asymptotic minimaxity and local asymptotic linearity, we study the asymptotic distribution of the following form

$$\Lambda^{(n)}_{f,\theta^{(n)}/\theta}(X^{(n)})=log\left(\frac{I_{\theta^{(n)}/f}(X^{(n)})}{I_{\theta/f}(X^{(n)})}\right),$$

where $I_{\theta^{(n)}/f}(X^{(n)})$ and $I_{\theta/f}(X^{(n)})$ are the conditional likelihood functions under $H_f^{(n)}(\theta^{(n)})$ and $H_f^{(n)}(\theta)$, respectively, such that

$$I_{\theta^{(n)}/f}(X^{(n)})=\prod_{i=1}^{s}\prod_{M=0}^{N-1}f(\sum_{j=0}^{i+sM-1}\pi_j(\theta_i^{(n)})X^{(n)}_{i+sM-j}),$$

and

$$I_{\theta/f}(X^{(n)})=\prod_{i=1}^{s}\prod_{M=0}^{N-1}f(\sum_{j=0}^{i+sM-1}\pi_j(\theta_i)X^{(n)}_{i+sM-j}).$$

Below we construct some random variables that will be needed for the sequel. We denote

- $Z_{i+sM}(\theta_i^{(n)})$ the residual under $H_f^{(n)}(\theta^{(n)})$ that coincides with $\varepsilon_{i+sM}(\theta_i^{(n)})$.
- $Z_{i+sM}(\theta_i)$ the residual under $H_f^{(n)}(\theta)$ that coincides with $\varepsilon_{i+sM}(\theta_i)$.

For $i=1,...,s$ we have

$$Z_{i+sM}(\theta_i)=\sum_{j=0}^{i+sM-1}\pi_j(\theta_i)X_{i+sM-j}+\sum_{k=0}^{\infty}\sum_{j=0}^{k}\pi_{j+i+sM}(\theta_i)\psi_{k-j}(\theta_i)Z_{-k}(\theta_i),$$

and

$$\gamma_{i,M,n}=\sum_{k=1}^{\infty}\sum_{j=1}^{k}\left[\pi_j(\theta_i^{(n)})-\pi_j(\theta_i)\right]\psi_{k-j}(\theta_i)Z_{i-k+sM}(\theta_i^{(n)}),$$

thus

$$Z_{i+sM}(\theta_i^{(n)}) = Z_{i+sM}(\theta_i) - \gamma_{i,M,n},$$

we use Taylor's expansion to expand the coefficient $\pi_j(\theta_i^{(n)})$ around θ_i, we obtain the following

$$\pi_j(\theta_i^{(n)})-\pi_j(\theta_i) = n^{-\frac{1}{2}}\alpha_i^{(n)}\frac{\partial\pi_j(\theta_i)}{\partial a_i}+n^{-\frac{1}{2}}\delta^{(n)}\frac{\partial\pi_j(\theta_i)}{\partial d}+\frac{n^{-1}}{2}\partial^2\pi_j(\theta_i+n^{-\frac{1}{2}}c\tau_i^{(n)}), \quad 0<c<1,$$

where

$$\partial^2\pi_j(\theta_i+n^{-\frac{1}{2}}c\tau_i^{(n)}) = \frac{\partial^2\pi_j(\theta_i+n^{-\frac{1}{2}}c\tau_i^{(n)})}{\partial a_i^2}(\alpha_i^{(n)})^2+\frac{\partial^2\pi_j(\theta_i+n^{-\frac{1}{2}}c\tau_i^{(n)})}{\partial d^2}(\delta^{(n)})^2$$
$$+\frac{\partial^2\pi_j(\theta_i+n^{-\frac{1}{2}}c\tau_i^{(n)})}{\partial a_i\partial d}\alpha_i^{(n)}\delta^{(n)}.$$

We pose

$$y_{i,M,n} = \frac{f^{\frac{1}{2}}(Z_{i+sM}-\gamma_{i,M,n})}{f^{\frac{1}{2}}(Z_{i+sM})}-1,$$

and

$$Z_{i,M,n} = \frac{n^{-\frac{1}{2}}}{2}\sum_{k=1}^{\infty}\sum_{j=1}^{k}(\frac{\partial}{\partial a_i}\pi_j(\theta_i)\alpha_i^{(n)}+\frac{\partial}{\partial d}\pi_j(\theta_i)\delta^{(n)})\psi_{k-j}(\theta_i)\phi_f(Z_{i+sM})Z_{i+sM-k},$$

when we replace $\frac{\partial}{\partial a_i}\pi_j(\theta_i)$, $\frac{\partial}{\partial d}\pi_j(\theta_i)$ and ψ_{k-j} by their values, we get

$$\begin{aligned} Z_{i,M,n} &= \frac{n^{-\frac{1}{2}}}{2}\sum_{k=1}^{\infty}((-d)\alpha_i^{(n)}+\frac{a_i}{k}\delta^{(n)})a_i^{k-1}\phi_f(Z_t)Z_{i+sM-k} \\ &= \frac{n^{-\frac{1}{2}}}{2}\sum_{k=1}^{q}((-d)\alpha_i^{(n)}+\frac{a_i}{k}\delta^{(n)})a_i^{k-1}\phi_f(Z_{i+sM})Z_{i+sM-k}+R_{i,M,n}, \end{aligned}$$

because, for all $i\in\{1,...,s\}$, we have

- $\frac{\partial}{\partial a_i}\pi_j(\theta_i) = j\frac{\Gamma(j-d)}{\Gamma(-d)\Gamma(j+1)}a_i^{j-1}$.
- $\frac{\partial}{\partial d}\pi_j(\theta_i) = \frac{\Gamma(j-d)}{\Gamma(-d)\Gamma(j+1)}a_i^j\sum_{h=1}^{j}\frac{1}{j-h-d}$.
- $\sum_{j=1}^{k}\frac{\Gamma(j-d)\Gamma(k-j+d)}{\Gamma(d)\Gamma(-d)\Gamma(k-j+1)\Gamma(j)} = -d$.

- $\displaystyle\sum_{j=1}^{k} \frac{\Gamma(j-d)\Gamma(k-j+d)}{\Gamma(d)\Gamma(-d)\Gamma(k-j+1)\Gamma(j+1)} \sum_{h=1}^{j} \frac{1}{j-h-d} = \frac{1}{k}, \quad \forall k \geq 1.$

and

$$R_{i,M,n} = \frac{n^{-\frac{1}{2}}}{2} \sum_{k=q+1}^{\infty} ((-d)\alpha_i^{(n)} + \frac{a_i}{k}\delta^{(n)})a_i^{k-1}\phi_f(Z_{i+sM})Z_{i+sM-k}.$$

For $i = 1, ..., s$, we write

$$\begin{aligned}
\sum_{M=0}^{N-1} Z_{i,M,n} &= \frac{n^{-\frac{1}{2}}}{2} \sum_{M=0}^{N-1} \sum_{k=1}^{q} ((-d)\alpha_i^{(n)} + \frac{a_i}{k}\delta^{(n)})a_i^{k-1}\phi_f(Z_{i+sM})Z_{i+sM-k} + \sum_{M=0}^{N-1} R_{i,M,n} \\
&= \frac{n^{-\frac{1}{2}}}{2} \sum_{k=1}^{q} \sum_{M=[\frac{k-i+1}{s}]}^{N-1} ((-d)\alpha_i^{(n)} + \frac{a_i}{k}\delta^{(n)})a_i^{k-1}\phi_f(Z_{i+sM})Z_{i+sM-k} + R_n,
\end{aligned}$$

$R_n = o_{p_{\theta_t}}(1)$ [94].

We consider the residual autocorrelation coefficients of order k denoted $r_{i,k,n}$ associated with f as follows

$$r_{i,k,n} = \frac{(\sigma^2 I(f))^{-\frac{1}{2}}}{N - [\frac{k-i+1}{s}]} \sum_{M=[\frac{k-i+1}{s}]}^{N-1} \phi_f(Z_{i+sM})Z_{i+sM-k},$$

we get

$$2\sum_{M=0}^{N-1} Z_{i,M,n} = \frac{n^{-\frac{1}{2}}}{(\sigma^2 I(f))^{-\frac{1}{2}}} \sum_{k=1}^{q} ((-d)\alpha_i^{(n)} + \frac{a_i}{k}\delta^{(n)})a_i^{k-1} \left(N - \left[\frac{k-i+1}{s}\right]\right) r_{i,k,n} + o_{p_\theta}(1).$$

For a period $s = 2$ (i.e., $i = 1$ or 2), we write

$$2\sum_{M=0}^{N-1} (Z_{1,M,n} + Z_{2,M,n}) = \frac{n^{-\frac{1}{2}}}{(\sigma^2 I(f))^{-\frac{1}{2}}} \sum_{k=1}^{q} ((-d)\alpha_1^{(n)} + \frac{a_1}{k}\delta^{(n)})a_1^{k-1} \left(N - \left[\frac{k}{2}\right]\right) r_{1,k,n} +$$

$$\frac{n^{-\frac{1}{2}}}{(\sigma^2 I(f))^{-\frac{1}{2}}} \sum_{k=1}^{q} ((-d)\alpha_2^{(n)} + \frac{a_2}{k}\delta^{(n)})a_2^{k-1} \left(N - \left[\frac{k-1}{2}\right]\right) r_{2,k,n} + o_{p_{\theta_t}}(1).$$

For any period $s > 2$, it would be better to use a matrix representation; for that we define a $(1 \times 2s)$-vector $\tau^{(n)} = (\underbrace{\alpha_1^{(n)}, ..., \alpha_s^{(n)}}_{s}, \underbrace{\delta^{(n)}, ..., \delta^{(n)}}_{s})$ and a $(2s \times 1)$-vector $\Delta_f^{(n)}$

as follows

$$\Delta_f^{(n)} = (n^{-\frac{1}{2}})(\sigma^2 I(f))^{\frac{1}{2}} \begin{pmatrix} -d\sum_{k=1}^{q} a_1^{k-1}\left(N - \left[\frac{k}{s}\right]\right) r_{1,k,n} \\ . \\ . \\ . \\ -d\sum_{k=1}^{q} a_s^{k-1}\left(N - \left[\frac{k-s+1}{s}\right]\right) r_{s,k,n} \\ \sum_{k=1}^{q} \frac{a_1^k}{k}\left(N - \left[\frac{k}{s}\right]\right) r_{1,k,n} \\ . \\ . \\ . \\ \sum_{k=1}^{q} \frac{a_s^k}{k}\left(N - \left[\frac{k-s+1}{s}\right]\right) r_{s,k,n} \end{pmatrix}, \tag{4.4}$$

finally, we find

$$2\sum_{i=1}^{s}\sum_{M=0}^{N-1} Z_{i,M,n} = \tau^{(n)}\Delta_f^{(n)} + o_{p_\theta}(1).$$

$\Delta_f^{(n)}$ is known as **central sequence**.

4.3 Main results

4.3.1 Local asymptotic normality

Proposition 4.1. *Assuming the regularity assumptions are valid, for all bounded sequences $\tau^{(n)}$ such that $sup||\tau^{(n)}|| < \infty$, we have*
i) The log likelihood ratio admits under $H_f^{(n)}(\theta)$ the asymptotic representation

$$\Lambda_{f,\theta^{(n)}/\theta}^{(n)}(X^{(n)}) = \tau^{(n)}\Delta_f^{(n)} - \frac{1}{2}(\tau^{(n)})\Gamma_f(\theta)(\tau^{(n)})^{\intercal} + o_{p_\theta}(1),$$

with $\Gamma_f(\theta) = I(f)\sigma^2\Gamma(\theta)$ and

$$\Gamma(\theta)=\begin{pmatrix}
\frac{d^2}{1-a_1^2} & 0 & \cdots & 0 & -\frac{d}{a_1}log(1-a_1^2) & 0 & \cdots & 0\\
0 & \frac{d^2}{1-a_2^2} & 0 & \vdots & 0 & -\frac{d}{a_2}log(1-a_2^2) & 0 & \vdots\\
\vdots & 0 & \ddots & 0 & \vdots & 0 & \vdots & 0\\
0 & \vdots & 0 & \frac{d^2}{1-a_s^2} & 0 & \vdots & 0 & -\frac{d}{a_s}log(1-a_s^2)\\
-\frac{d}{a_1}log(1-a_1^2) & 0 & \vdots & 0 & \sum_{k=1}^{\infty}(\frac{a_1^k}{k})^2 & 0 & \vdots & 0\\
0 & -\frac{d}{a_2}log(1-a_2^2) & 0 & \vdots & 0 & \sum_{k=1}^{\infty}(\frac{a_2^k}{k})^2 & 0 & \vdots\\
\vdots & 0 & \ddots & 0 & \vdots & 0 & \ddots & 0\\
0 & \cdots & 0 & \frac{d^2}{1-a_s^2} & 0 & \cdots & 0 & \sum_{k=1}^{\infty}(\frac{a_s^k}{k})^2,
\end{pmatrix},$$

ii) Still under $H_f^{(n)}(\theta)$, as $n \longrightarrow \infty$, $\Delta_f^{(n)}$is asymptotically normal with mean 0 and covariance equal to $\Gamma_f(\theta)$.

The proof of the aforementioned proposition is based on a slightly modified version of Swensen's lemma in accordance with the periodic models; this modified form was considered in several works, such as [27].

To proceed with the proof, we must examine the following conditions:

- (C1) $\mathbb{E}\sum_{M=0}^{N-1}\sum_{i=1}^{s}(Z_{i,M,n}-y_{i,M,n})^2 \underset{N\to\infty}{\longrightarrow} 0.$
- (C2) $\sup_N \sum_{M=0}^{N-1}\sum_{i=1}^{s}\mathbb{E}(Z_{i,M,n})^2 < \infty.$
- (C3) $\max_{1\leq i\leq s}\max_{0\leq M\leq N-1}|Z_{i,M,n}| \underset{N\to\infty}{\longrightarrow} 0$ under $\mathbb{P}_{\theta_t}$.
- (C4) $\sum_{M=0}^{N-1}\sum_{i=1}^{s}Z_{i,M,n}^2-\frac{(\lambda^{(n)})^2}{4} \underset{N\to\infty}{\longrightarrow} 0$ under $\mathbb{P}_{\theta_t}$.
- (C5) $\sum_{M=0}^{N-1}\sum_{i=1}^{s}\mathbb{E}(Z_{i,M,n}^2 I_{(|Z_{i,M,n}|>\frac{1}{2})}|A_{n,i+sM-1}) \underset{N\to\infty}{\longrightarrow} 0$ under $\mathbb{P}_{\theta_t}$.
- (C6) $\mathbb{E}(Z_{i,M,n}|A_{n,i+sM-1})=0 \quad$ for $\quad i=1,...,s.$

If the six conditions (C1)-(C6) hold, then

$$\Lambda_{f,\theta^{(n)}/\theta}^{(n)}=2\sum_{M=0}^{N-1}\sum_{i=1}^{s}Z_{i,M,n}-\frac{(\lambda^{(n)})^2}{2}+o_{p_\theta}(1), \tag{4.5}$$

and $2\sum_{M=0}^{N-1}\sum_{i=1}^{s} Z_{i,M,n}$ is asymptotically normal with mean 0 and variance equal to $(\lambda^{(n)})^2$.

Remark 4.1. The decomposition (4.5) is called the **Local Asymptotic Quadratic (LAQ)** decomposition.

Proof 15 *The conditions (C1)-(C6) will be checked one by one.*

(C1) $\mathbb{E}\sum_{M=0}^{N-1}\sum_{i=1}^{s}(Z_{i,M,n}-y_{i,M,n})^2 \underset{N\to\infty}{\longrightarrow} 0.$

$$(\frac{f^{\frac{1}{2}}(Z_{i+sM}-\gamma_{i,M,n})}{f^{\frac{1}{2}}(Z_{i+sM})}-1-Z_{i,M,n})^2 \leq 2\left(\frac{f^{\frac{1}{2}}(Z_{i+sM}-\gamma_{i,M,n})}{f^{\frac{1}{2}}(Z_{i+sM})}-1-\frac{1}{2}\gamma_{i,M,n}\phi_f(Z_{i+sM})\right)^2$$
$$+2\left(Z_{i,M,n}-\frac{1}{2}\gamma_{i,M,n}\phi_f(Z_{i+sM})\right)^2.$$

We begin with the second term $\left(Z_{i,M,n}-\frac{1}{2}\gamma_{i,M,n}\phi_f(Z_{i+sM})\right)^2$*, such that*

$$\begin{aligned}
\gamma_{i,M,n}\phi_f(Z_{i+sM}) &= \sum_{k=1}^{\infty}\sum_{j=1}^{k}\left[\pi_j(\theta_i^{(n)})-\pi_j(\theta_i)\right]\psi_{k-j}(\theta_i)Z_{i-k+sM}\phi_f(Z_{i+sM})\\
&= \sum_{k=1}^{\infty}\sum_{j=1}^{k}[n^{-\frac{1}{2}}\alpha_i^{(n)}\frac{\partial\pi_j(\theta_i)}{\partial a_i}+n^{-\frac{1}{2}}\delta^{(n)}\frac{\partial\pi_j(\theta_i)}{\partial d}+\frac{n^{-1}}{2}\partial^2\pi_j(\theta_i+n^{-\frac{1}{2}}c\tau_i^{(n)})]\\
&\times\ \psi_{k-j}(\theta_i)Z_{i-k+sM}\phi_f(Z_{i+sM})\\
&= 2Z_{i,M,n}+\sum_{k=1}^{\infty}\sum_{j=1}^{k}(\frac{n^{-1}}{2}\partial^2\pi_j(\theta_i+n^{-\frac{1}{2}}c\tau_i^{(n)}))\psi_{k-j}(\theta_i)Z_{i-k+sM}\phi_f(Z_{i+sM}).
\end{aligned}$$

We conduct a simple calculation and we find

$$\begin{aligned}
(Z_{i,M,n}-\frac{1}{2}\gamma_{i,M,n}\phi_f(Z_{i+sM}))^2 &= [(D_a+(D_d)+(D_{ad})]^2\\
&\leq 5\left[(D_a)^2+(D_d)^2+(D_{ad})^2\right],
\end{aligned}$$

$$\mathbb{E}\sum_{M=0}^{N-1}\sum_{i=1}^{s}(Z_{i,M,n}-\frac{1}{2}\gamma_{i,M,n}\phi_f(Z_{i+sM}))^2 \leq 5\sum_{M=0}^{N-1}\sum_{i=1}^{s}\left[\mathbb{E}(D_a)^2+\mathbb{E}(D_d)^2+\mathbb{E}(D_{ad})^2\right],$$

where

$$\begin{aligned}
D_a &= \frac{-n^{-1}}{4}\sum_{k=1}^{\infty}\sum_{j=1}^{k}\frac{\partial^2\pi_j(\theta_i+n^{-\frac{1}{2}}c\tau_i^{(n)})}{\partial a_i^2}(\alpha_i^{(n)})^2\psi_{k-j}(\theta_i)Z_{i-k+sM}(\theta_i^{(n)})\phi_f(Z_{i+sM})\\
&= \frac{-n^{-1}}{4}(\alpha_i^{(n)})^2\left[(1-(a_i+n^{-\frac{1}{2}}c\alpha_i^{(n)})L)^{d+n^{-\frac{1}{2}}c\delta^{(n)}-2}(1-a_iL)^{-d}Z_{i+sM-2}\right]\phi_f(Z_{i+sM}),
\end{aligned}$$

the terms $(1-(a_i+n^{-\frac{1}{2}}c\alpha_i^{(n)})L)^{d+n^{-\frac{1}{2}}c\delta^{(n)}-2}$ *and* $(1-a_iL)^{-d}$ *are expressed in terms of summable coefficients, thus of square summable coefficients, we deduce that* $\sum_{M=0}^{N-1}\sum_{i=1}^{s}\mathbb{E}(D_a)^2$ *is finite and tends to* 0 *as* N *tends to infinity.*

$$\begin{aligned}
D_d &= \frac{-n^{-1}}{4}\sum_{k=1}^{\infty}\sum_{j=1}^{k}\frac{\partial^2\pi_j(\theta_i+n^{-\frac{1}{2}}c\tau_i^{(n)})}{\partial d^2}(\delta^{(n)})^2\psi_{k-j}(\theta_i)Z_{i-k+sM}(\theta_i^{(n)})\phi_f(Z_{i+sM})\\
&= \frac{-n^{-1}}{4}(\delta^{(n)})^2[log(1-(a_i+n^{-\frac{1}{2}}c\alpha_i^{(n)})L)^2(1-(a_i+n^{-\frac{1}{2}}c\alpha_i^{(n)})L)^{d+n^{-\frac{1}{2}}c\delta^{(n)}}\\
\times &\ (1-a_iL)^{-d}Z_{i+sM}]\phi_f(Z_{i+sM}),
\end{aligned}$$

the operator that appears in the final form is of a summable coefficient, as it is expressed as a product of three summable coefficients. Thus $\sum_{M=0}^{N-1}\sum_{i=1}^{s}\mathbb{E}(D_d)^2 \underset{N\to\infty}{\longrightarrow} 0.$

$$\begin{aligned}
D_{ad} &= \frac{-n^{-1}}{2}\sum_{k=1}^{\infty}\sum_{j=1}^{k}\frac{\partial^2\pi_j(\theta_i+n^{(-1/2)}c\tau_i^{(n)})}{\partial a_i\partial d}\alpha_i^{(n)}\delta^{(n)}\psi_{k-j}(\theta_i)Z_{i-k+sM}(\theta_i^{(n)})\phi_f(Z_{i+sM})\\
&= \frac{-n^{-1}}{2}\alpha_i^{(n)}\delta^{(n)}[log(1-(a_i+n^{-\frac{1}{2}}c\alpha_i^{(n)})L)(1-(a_i+n^{-\frac{1}{2}}c\alpha_i^{(n)})L)^{d+n^{-\frac{1}{2}}c\delta^{(n)}-1}\\
\times &\ (1-a_iL)^{-d}Z_{i+sM}]\phi_f(Z_{i+sM}),
\end{aligned}$$

the final form has an operator expressed as a product of three terms of square summable coefficients, so $\sum_{M=0}^{N-1}\sum_{i=1}^{s}\mathbb{E}(D_{ad})^2$ *tends to* 0 *as* N *tends to infinity.*

Now we deal with the other term $\left(\frac{f^{\frac{1}{2}}(Z_{i+sM}-\gamma_{i,M,n})}{f^{\frac{1}{2}}(Z_{i+sM})}-1-\frac{1}{2}\gamma_{i,M,n}\phi_f(Z_{i+sM})\right)^2.$

From the previous calculation of $\gamma_{i,M,n}$*, we can write*

$$\begin{aligned}
\gamma_{i,M,n} &= (n^{-\frac{1}{2}}\sum_{k=1}^{\infty}(-d\alpha_i^{(n)}+\frac{a_i}{k}\delta^{(n)})a_i^{k-1}+\sum_{k=1}^{\infty}\sum_{j=1}^{k}\frac{n^{-1}}{2}\partial^2\pi_j(\theta_i+n^{-\frac{1}{2}}c\tau_i^{(n)})\psi_{k-j}(\theta_i))Z_{i-k+pM}\\
&= n^{-\frac{1}{2}}U_{i,M,n}+R_{i,M,n}+V_{i,M,n},
\end{aligned}$$

where we have denoted the following

$$Q_{1,k}^{(n)} = (-d\alpha_i^{(n)}+\frac{a_i}{k}\delta^{(n)})a_i^{(k-1)},$$

$$Q_{2,k}^{(n)}(\theta_i+n^{-\frac{1}{2}}c\tau_i^{(n)}) = \sum_{j=1}^{k}\psi_{k-j}(\theta_i)\partial^2\pi_j(\theta_i+n^{-\frac{1}{2}}c\tau_i^{(n)}),$$

and

$$U_{i,M,n} = \sum_{k=1}^{q} Q_{1,k}^{(n)} Z_{i-k+sM},$$

$$R_{i,M,n} = n^{-\frac{1}{2}} \sum_{k=q+1}^{\infty} Q_{1,k}^{(n)} Z_{i-k+sM},$$

$$V_{i,M,n} = \frac{n^{-1}}{2} \sum_{k=1}^{\infty} Q_{2,k}(\theta_i + n^{-\frac{1}{2}} c\tau_i^{(n)}) Z_{i-k+sM},$$

q is the truncation parameter. We deal with these last three variables.
i)

$$\begin{aligned}\mathbb{E}(|V_{i,M,n}|^2) &= \frac{n^{-2}}{4}\mathbb{E}(\sum_{k=1}^{\infty} Q_{2,k}(\theta_i + n^{-\frac{1}{2}} c\tau_i^{(n)}) Z_{i-k+sM})^2 \\ &\leq 5\frac{n^{-2}}{4}\mathbb{E}((D_a)^2 + (D_d)^2 + (D_{ad})^2),\end{aligned}$$

and consequently $V_{i,M,n} \underset{N\to\infty}{\longrightarrow} 0$ *(see [94] for more details).*
ii)

$$R_{i,M,n} = n^{-\frac{1}{2}} \sum_{k=q+1}^{\infty} Q_{1,k}^{(n)} Z_{i-k+sM},$$

$\mathbb{E}(R_{i,M,n}^4) = n^{-2}O(1)$ *when* $N \longrightarrow \infty$ *and then* $R_{i,M,n} \longrightarrow 0$ *[94].*
We denote

$$W_{i,M,n} = V_{i,M,n} + R_{i,M,n},$$

then

$$\gamma_{i,M,n} = n^{-\frac{1}{2}} U_{i,M,n} + W_{i,M,n},$$

we consider

$$B_i(k) = \{|Z_{i+sM-l}| < k, l = 1, ..., q\}; \quad for \quad i = 1, ..., s.$$

$$C_i(k) = \{|Z_{i+sM-l}| \geq k, l = 1, ..., q\}; \quad for \quad i = 1, ..., s.$$

Let

$$\begin{aligned}
B_{i,n} &= \sum_{M=0}^{N-1}\mathbb{E}\left\{\mathbb{1}_{B_i(k)}\left(\frac{f^{\frac{1}{2}}(Z_{i+sM}-\gamma_{i,M,n})}{f^{\frac{1}{2}}(Z_{i+sM})}-1-\frac{1}{2}\gamma_{i,M,n}\phi_f(Z_{i+sM})\right)^2\right\}\\
&= \sum_{M=0}^{N-1}\mathbb{E}\left\{\mathbb{1}_{B_i(k)}\left(\frac{f^{\frac{1}{2}}(Z_{i+sM}-n^{-\frac{1}{2}}U_{i,M,n}-W_{i,M,n})-f^{\frac{1}{2}}(Z_{i+sM})}{(n^{-\frac{1}{2}}U_{i,M,n}+W_{i,M,n})f^{\frac{1}{2}}(Z_{i+sM})}-\frac{1}{2}\phi_f(Z_{i+sM})\right)^2\gamma^2_{i,M,n}\right\}\\
&= \sum_{M=0}^{N-1}\iint_{B_i(k)}(n^{-\frac{1}{2}}u+w)^2\int_{\mathbb{R}}\left(\frac{f^{\frac{1}{2}}(z-n^{-\frac{1}{2}}u-w)-f^{\frac{1}{2}}(z)}{(n^{-\frac{1}{2}}u+w)f^{\frac{1}{2}}(z)}-\frac{1}{2}\frac{f'(z)}{f(z)}\right)^2 dzg(u,w)dudw\\
&\leq \sum_{M=0}^{N-1}\sup_{\substack{|u|<k_0\\|w|<n^{-r}k_1}}\int_{\mathbb{R}}\left(\frac{f^{\frac{1}{2}}(z-n^{-\frac{1}{2}}u-w)-f^{\frac{1}{2}}(z)}{(n^{-\frac{1}{2}}u+w)}-\frac{1}{2}\frac{f'(z)}{f^{\frac{1}{2}}(z)}\right)^2 dz\iint_{\substack{|u|<k_0\\|w|<n^{-r}k_1}}(n^{-\frac{1}{2}}u+w)^2g(u,w)d
\\
&\leq \sup_{\substack{|u|<k_0\\|w|<n^{-r}k_1}}\int_{\mathbb{R}}\left(\frac{f^{\frac{1}{2}}(z-n^{-\frac{1}{2}}u-w)-f^{\frac{1}{2}}(z)}{(n^{-\frac{1}{2}}u+w)}-\frac{1}{2}\frac{f'(z)}{f^{\frac{1}{2}}(z)}\right)^2 dz\sum_{M=0}^{N-1}\mathbb{E}(\gamma^2_{i,M,n}),
\end{aligned}$$

we put

$$C_n(K_0,K_1)=\sup_{\substack{|u|<k_0\\|w|<n^{-r}k_1}}\int_{\mathbb{R}}\left(\frac{f^{\frac{1}{2}}(z-n^{-\frac{1}{2}}u-w)-f^{\frac{1}{2}}(z)}{(n^{-\frac{1}{2}}u+w)}-\frac{1}{2}\frac{f'(z)}{f^{\frac{1}{2}}(z)}\right)dz,$$

according to [98] (lemma 2) $C_n(K_0,K_1)\underset{n\to\infty}{\longrightarrow}0$.

It remains to verify the convergence of $\sum_{M=0}^{N-1}\mathbb{E}(\gamma^2_{i,M,n})$, *we have*

$$\mathbb{E}(\gamma^2_{i,M,n})\leq 2n^{-1}\sum_{k=0}^{\infty}\sigma^2(Q^{(n)}_{1,k})^2+\frac{n^{-1}}{4}\sum_{k=0}^{\infty}\sigma^2(Q^{(n)}_{2,k})^2<\infty,$$

then

$$\sum_{M=0}^{N-1}\mathbb{E}(\gamma^2_{i,M,n})\leq\frac{2N}{n}\sigma^2\sum_{k=0}^{\infty}(Q^{(n)}_{1,k})^2+\frac{N}{4n}\sigma^2\sum_{k=0}^{\infty}(Q^{(n)}_{2,k})^2<\infty,$$

it results that

$$B_{i,n}\leq C_n(K_0,K_1)\sum_{M=0}^{N-1}\mathbb{E}(\gamma^2_{i,M,n})\underset{n\to\infty}{\longrightarrow}0.$$

Same as $B_{i,n}$, we deal with $C_{i,n}$

$$\begin{aligned}
C_{i,n} &= \sum_{M=0}^{N-1}\mathbb{E}\left\{\mathbb{1}_{C_i(k)}\left(\frac{f^{\frac{1}{2}}(Z_{i+sM}-n^{-\frac{1}{2}}U_{i,M,n}-W_{i,M,n})-f^{\frac{1}{2}}(Z_{i+sM})}{f^{\frac{1}{2}}(Z_{i+sM})}-\frac{1}{2}\gamma_{i,M,n}\phi_f(Z_{i+sM})\right)^2\right\}\\
&= \sum_{M=0}^{N-1}\iint_{C_i(k)} dg_i(u,v)\int_{\mathbb{R}}(\left(f^{-\frac{1}{2}}(z-n^{-\frac{1}{2}}u-w)f^{-\frac{1}{2}}-\frac{1}{2}(n^{-\frac{1}{2}}u+w)\phi_f(Z_{i+sM})f^{-\frac{1}{2}}\right)^2 dzdudw\\
&\le I(f)\sum_{M=0}^{N-1}\mathbb{E}\left\{\mathbb{1}_{C_i(k)}\quad \gamma_{i,M,n}^2\right\}\\
&\le I(f)\sum_{M=0}^{N-1}\mathbb{E}\left\{\mathbb{1}_{\{Z_{i+sM-l}\ge k;l=1,\dots,q\}}\quad \gamma_{i,M,n}^2\right\},
\end{aligned}$$

using Holder's inequality, we obtain

$$C_{i,n}\le I(f)\mathbb{P}\left(\{|Z_{i+sM-l}|\ge k;l=1,...,q\}\right)^{\frac{1}{3}}\sum_{M=0}^{N-1}\mathbb{E}[|\gamma_{i,M,n}|^3]^{\frac{2}{3}}.$$

Now we simply need to show that $\sum_{M=0}^{N-1}\mathbb{E}[|\gamma_{i,M,n}|^3]^{\frac{2}{3}}$ *is bounded, we have*

$$\mathbb{E}(\gamma_{i,M,n}^4)\le 4n^{-2}\mathbb{E}(\sum_{k=1}^{\infty}Q_{1,k}Z_{i-k+sM})^4+\frac{n^{-4}}{2}\mathbb{E}(\sum_{k=1}^{\infty}Q_{2,k}(\theta_i+n^{-\frac{1}{2}}\delta^{(n)})Z_{i-k+sM})^4=O(n^{-2}),$$

the coefficients $Q_{1,k}$ and $Q_{2,k}(\theta_i+n^{-\frac{1}{2}}\delta^{(n)})$ are square summable, thus $\mathbb{E}[|\gamma_{i,M,n}|^3]^{\frac{2}{3}}\le \mathbb{E}(|\gamma_{i,M,n}|^4)^{\frac{1}{2}}=O(n^{-1})$ as $n\longrightarrow\infty$. Then $\mathbb{E}[|\gamma_{i,M,n}|^3]^{\frac{2}{3}}$ is bounded.
Moreover, it is enough to select a large k, such as $\mathbb{P}\left(|Z_{i+sM-l}|\ge k;l=1,...,q\right)$ be as small as we need.
Let $k'=k(qlogq)^{\frac{1}{2}}$.

$$\mathbb{P}\left(|Z_{i+sM-l}|\ge k';l=1,...,q\right)\le\sum_{l=1}^{q}\mathbb{P}\left(|Z_{i+sM-l}|\ge k'\right)\le\sum_{l=1}^{q}\sigma^2(k')^{-2}=\sigma^2k^{-2}(logq)^{-1},$$

$$\left[\mathbb{P}\left(|Z_{i+sM-l}|\ge k;l=1,...,q\right)\right]^{\frac{1}{3}}\le\sigma^{\frac{2}{3}}k^{\frac{-2}{3}}(\log q)^{\frac{-1}{3}},$$

then $C_{i,n}=o(1)$ as N tends to infinity. Finally, it is easy to verify that

$$\mathbb{E}\sum_{M=0}^{N-1}\sum_{i=1}^{s}(Z_{i,M,n}-y_{i,M,n})^2\le 10\sum_{M=0}^{N-1}\sum_{i=1}^{s}\left[\mathbb{E}(D_a)^2+\mathbb{E}(D_d)^2+\mathbb{E}(D_{ad})^2\right]+2\sum_{M=0}^{N-1}\sum_{i=1}^{s}(B_{i,n}+C_{i,n})\underset{N\to\infty}{\longrightarrow}0.$$

(C2) $\sup_N\sum_{M=0}^{N-1}\sum_{i=1}^{s}\mathbb{E}(Z_{i,M,n})^2<\infty.$

$$Z_{i,M,n}=\frac{n^{-\frac{1}{2}}}{2}\sum_{k=1}^{\infty}Q_{1,k}^{(n)}\phi_f(Z_{i+sM})Z_{i-k+sM},$$

$$\begin{aligned}\mathbb{E}(Z_{i,M,n}^2) &\leq \frac{n^{-1}}{4}\lim_{m\to\infty}\inf \mathbb{E}\left[\sum_{k=1}^{m} Q_{1,k}^{(n)}\phi_f(Z_{i+sM})Z_{i-k+sM}\right]^2 \\ &= \frac{n^{-1}}{4}\lim_{m\to\infty}\inf\sum_{k=1}^{m}(Q_{1,k}^{(n)})^2 I(f)\sigma^2,\end{aligned}$$

$$\sum_{M=0}^{N-1}\mathbb{E}(Z_{i,M,n}^2) \leq \frac{Nn^{-1}}{4}I(f)\sigma^2\sum_{k=1}^{\infty}(Q_{1,k}^{(n)})^2,.$$

since $\sum_{k=1}^{\infty}(Q_{1,k}^{(n)})^2<\infty$, *we conclude that* $\sup_N \sum_{M=0}^{N-1}\sum_{i=1}^{s}\mathbb{E}(Z_{i,M,n})^2<\infty.$

(C3) $\max_{1\leq i\leq s}\max_{0\leq M\leq N-1}|Z_{i,M,n}| \underset{N\to\infty}{\longrightarrow} 0$ *under* $\mathbb{P}_{\theta_t}$.

For a fixed $i=1,...,s$ *by Chebyshev's inequality we have*

$$\begin{aligned}\mathbb{P}\left(\max_{0\leq M\leq N-1}|Z_{i,M,n}|>\varepsilon\right) &= \mathbb{P}\left(\sum_{M=0}^{N-1}Z_{i,M,n}^2 I_{\{|Z_{i,M,n}|>\varepsilon\}}>\varepsilon^2\right) \\ &\leq \frac{1}{\varepsilon^2}\mathbb{E}\left(\sum_{M=0}^{N-1}Z_{i,M,n}^2 I_{\{|Z_{i,M,n}|>\varepsilon\}}\right) \\ &\leq \frac{1}{\varepsilon^2}\sum_{M=0}^{N-1}\mathbb{E}\left(Z_{i,M,n}^2 I_{\{|Z_{i,M,n}|>\varepsilon\}}\right),\end{aligned}$$

by Fatou's lemma we write

$$\begin{aligned}\sum_{M=0}^{N-1}\mathbb{E}\left(Z_{i,M,n}^2 I_{\{|Z_{i,M,n}|>\varepsilon\}}\right) &\leq \frac{1}{4n}\lim_{m\to\infty}\inf\sum_{M=0}^{N-1}\mathbb{E}\left(\left[\sum_{k=1}^{m}Q_{1,k}\phi_f(Z_{i+sM})Z_{i+sM-k}\right]^2 I_{\{|Z_{i,M,n}|>\varepsilon\}}\right) \\ &\leq \frac{1}{4n}\lim_{m\to\infty}\inf\sum_{M=0}^{N-1}\int_{\{|Z_{i,M,n}|>\varepsilon\}}\left(\sum_{k=1}^{m}Q_{1,k}\phi_f(Z_{i+sM})Z_{i+sM-k}\right)^2 d\mathbb{P}_{\theta_t},\end{aligned}$$

we have

$$\int_{\{|Z_{i,M,n}|>\varepsilon\}}\left(\sum_{k=1}^{m}Q_{1,k}\phi_f(Z_{i+sM})Z_{i+sM-k}\right)^2 d\mathbb{P}_{\theta_t} \leq \sum_{k=1}^{m}Q_{1,k}^2\int_{\{|Z_{i,M,n}|>\varepsilon\}}\left(\phi_f(Z_{i+sM})Z_{i+sM-k}\right)^2 d\mathbb{P}_{\theta_t},$$

but if

$$|Z_{i,M,n}|>\varepsilon \Leftrightarrow \left|\frac{n^{-\frac{1}{2}}}{2}\sum_{k=1}^{\infty}Q_{1,k}\phi_f(Z_{i+sM})Z_{i+sM-k}\right|>\varepsilon,$$

then

$$\sum_{M=0}^{N-1}\int_{\{|Z_{i,M,n}|>\varepsilon\}}\left(\sum_{k=1}^{m}Q_{1,k}\phi_f(Z_{i+sM})Z_{i+sM-k}\right)^2 d\mathbb{P}_{\theta_t}$$

$$\leq \frac{1}{n}\sum_{M=0}^{N-1} \int\limits_{\left\{\left|Z_{i,M,n}\right|>\varepsilon\right\}} \left(\sum_{k=1}^{m} Q_{1,k}\phi_f(Z_{i+sM})Z_{i+sM-k}\right)^2 d\mathbb{P}_{\theta_t}$$

$$\leq \frac{1}{n}\sum_{k=1}^{m} Q_{1,k}\sum_{M=0}^{N-1} \int\limits_{\left\{\left|U_{i,M,n}\right|>n^{\frac{1}{2}}\varepsilon\right\}} \left(\phi_f(Z_{i+sM})Z_{i+sM-k}\right)^2 d\mathbb{P}_{\theta_t}$$

because

$$\int\limits_{\left\{\left|U_{i,M,n}\right|>n^{\frac{1}{2}}\varepsilon\right\}} \left(\phi_f(Z_{i+sM})Z_{i+sM-k}\right)^2 d\mathbb{P}_{\theta_t} \underset{n\to\infty}{\longrightarrow} 0.$$

Thus $\max\limits_{1\leq i\leq s}\max\limits_{0\leq M\leq N-1}|Z_{i,M,n}| \underset{N\to\infty}{\longrightarrow} 0$ *under* $\mathbb{P}_{\theta_t}$.

(C4) $\sum\limits_{M=0}^{N-1}\sum\limits_{i=1}^{s} Z_{i,M,n}^2 - \frac{(\lambda^{(n)})^2}{4} \underset{N\to\infty}{\longrightarrow} 0$ *under* $\mathbb{P}_{\theta_t}$.

We have

$$\sum_{M=0}^{N-1}\sum_{i=1}^{s} Z_{i,M,n}^2 = \frac{n^{-1}}{4}\sum_{M=0}^{N-1}\sum_{i=1}^{s}\left(\sum_{k=1}^{\infty} Q_{1,k}^{(n)}\phi_f(Z_{i+sM})Z_{i-k+sM}\right)^2,$$

we set for $i = 1, ..., s$

$$U_{i,M,n} = \frac{n^{-1}}{4}\sum_{M=0}^{N-1}\phi_f^2(Z_{i+sM})\left(\sum_{k=1}^{m} Q_{1,k}^{(n)} Z_{i-k+sM}\right)^2,$$

$$y_{i,M,n} = \phi_f^2(Z_{i+sM})\left(\sum_{k=1}^{m} Q_{1,k}^{(n)} Z_{i-k+sM}\right)^2,$$

the process $(y_{i,M,n})$ *is ergodic with mean* $I(f)\sigma^2\sum\limits_{k=1}^{m}(Q_{1,k}^{(n)})^2$, *so* $U_{i,M,n}$ *is such that*

$$U_{i,M,n} - \frac{1}{4}I(f)\sigma^2\sum_{k=1}^{m}(Q_{1,k}^{(n)})^2 = o_{p_{\theta_t}}(1).$$

And $\frac{1}{4}I(f)\sigma^2\sum\limits_{k=1}^{m}(Q_{1,k}^{(n)})^2 \underset{m\to\infty}{\longrightarrow} \frac{(\lambda_i^{(n)})^2}{4} = \frac{1}{4}I(f)\sigma^2\sum\limits_{k=1}^{\infty}(Q_{1,k}^{(n)})^2 < \infty$.

On the other hand, we apply the Cauchy-Schwartz inequality, and we obtain

$$\limsup_N \mathbb{P}\left\{\left|\sum_{M=0}^{N-1}\sum_{i=1}^{s} Z_{i,M,n}^2 - U_{i,M,n}\right| > \varepsilon\right\} \leq$$

$$\frac{1}{4}I(f)\left\{\sum_{k=m+1}^{\infty} Q_{1,k}^2\sigma^2 + 2\left[\sum_{k=1}^{m} Q_{1,k}^2\sigma^2\right]^{\frac{1}{2}}\left[\sum_{k=m+1}^{\infty} Q_{1,k}^2\sigma^2\right]^{\frac{1}{2}}\right\} \underset{m\to\infty}{\longrightarrow} 0,$$

thus according to [11]

$$U_{i,M,n} - \frac{(\lambda_i^{(n)})^2}{4} = o_{p_{\theta_t}}(1),$$

where $(\lambda^{(n)})^2 = \sum_{i=1}^{s}(\lambda_i^{(n)})^2$.

(C5) $\sum_{M=0}^{N-1}\sum_{i=1}^{s}\mathbb{E}(Z_{i,M,n}^2 I_{(|Z_{i,M,n}|>\frac{1}{2})}|A_{n,i+sM-1}) \underset{N\to\infty}{\longrightarrow} 0$ *under* $\mathbb{P}_{\theta_t}$.

Note that

$$\begin{aligned}\mathbb{E}\left(\mathbb{E}\left(Z_{i,M,n}^2 I_{(|Z_{i,M,n}|>\frac{1}{2})}|A_{i+sM-1}\right)\right) &= \mathbb{E}\left(Z_{i,M,n}^2 I_{(|Z_{i,M,n}|>\frac{1}{2})}\right)\\ &\leq \mathbb{E}\left(Z_{i,M,n}^2 I_{\left(\max\limits_{1\leq i\leq s}\max\limits_{0\leq M\leq N-1}|Z_{i,M,n}|>\frac{1}{2}\right)}\right),\end{aligned}$$

following Swensen's lemma (1985, p67) [98], it suffices to show that $\sum_{M=0}^{N-1}\sum_{i=1}^{s}Z_{i,M,n}^2$ *is uniformly integrable. To do this, we apply Serfling's lemma. And by condition* $(C4)$, *under* $\mathbb{P}_{\theta_t}$ *we have*

$$\sum_{M=0}^{N-1}\sum_{i=1}^{s}Z_{i,M,n}^2 - \frac{\left(\lambda^{(n)}\right)^2}{4} = o_{p_\theta}(1), \quad as \quad n\to\infty,$$

and

$$\sum_{M=0}^{N-1}\sum_{i=1}^{s}\mathbb{E}\left(Z_{i,M,n}^2\right) - \frac{\left(\lambda^{(n)}\right)^2}{4} = o(1), \quad as \quad n\to\infty.$$

Consequently $\sum_{M=0}^{N-1}\sum_{i=1}^{s}Z_{i,M,n}^2$ *is uniformly integrable.*

(C6) $\mathbb{E}(Z_{i,M,n}|A_{n,i+sM-1}) = 0 \quad$ *for* $\quad i = 1, ..., s$.

We have

$$\begin{aligned}\mathbb{E}\left(Z_{i,M,n}|A_{i+sM-1}\right) &= \frac{n^{-\frac{1}{2}}}{2}\sum_{k=1}^{n-1}Q_{1,k}\mathbb{E}\left(Z_{i-k+sM}\phi\left(Z_{i+sM}\right)|A_{i+sM-1}\right)\\ &= \frac{n^{-\frac{1}{2}}}{2}\sum_{k=1}^{n-1}Q_{1,k}Z_{i-k+sM}\mathbb{E}\left(\phi\left(Z_{i+sM}\right)|A_{i+sM-1}\right),\end{aligned}$$

and

$$\mathbb{E}\left(\phi_f(Z_{i+sM})\right) = \int_R -\frac{f'(z)}{f(z)}f(z)dz = -\int_R f'(z)dz = 0.$$

Because if we set $z = F^{-1}(u) = inf\{x/f(x) \leq u\}$ and $u \in [0,\ 1]$ (F is the distribution function of f of the process $(\varepsilon_t, t \in Z)$), we write $dz = dF^{-1}(u) = \dfrac{du}{f(F^{-1}(u))}$. By Hàjek and Sidak (1967) [43] we have

$$\mathbb{E}\left(\phi_f(Z_{i+sM})\right) = \int_0^1 -\frac{f'(F^{-1}(u))}{f(F^{-1}(u))} du = 0.$$

Remark 4.2. According to condition $(C4)$, $(\lambda^{(n)})^2 = I(f)\sigma^2 \sum_{k=1}^{\infty}\sum_{i=1}^{s} Q_{1,k}^2$.
Such that $Q_{1,k}^2 = (-d\alpha_i^{(n)} + \frac{a_i}{k}\delta^{(n)})^2 a_i^{2(k-1)}$.

$$\begin{aligned}
\sum_{k=1}^{\infty}\sum_{i=1}^{s} Q_{1,k}^2 &= \sum_{k=1}^{\infty}\sum_{i=1}^{s}(-d\alpha_i^{(n)} + \frac{a_i}{k}\delta^{(n)})^2 a_i^{2(k-1)} \\
&= \sum_{i=1}^{s}(d\alpha_i^{(n)})^2 \sum_{i=1}^{s}\sum_{k=1}^{\infty} a_i^{2(k-1)} - 2d\delta^{(n)} \sum_{i=1}^{s}\sum_{k=1}^{\infty} \alpha_i \frac{a_i^{2(k-1)}}{k} + (\delta^{(n)})^2 \sum_{i=1}^{s}\sum_{k=1}^{\infty} \frac{a_i^{2k}}{k^2} \\
&= \sum_{i=1}^{s} \frac{(d\alpha_i^{(n)})^2}{1-a_i^2} - 2d\delta^{(n)} \sum_{i=1}^{s} \frac{\alpha_i}{a_i} log(1-a_i^2) + (\delta^{(n)})^2 \sum_{i=1}^{s}\sum_{k=1}^{\infty} (\frac{a_i^k}{k})^2.
\end{aligned}$$

For $s = 2$. We write

$$\begin{gathered}
(\lambda^{(n)})^2 = I(f)\sigma^2[d^2(\frac{(\alpha_1^{(n)})^2}{1-a_1^2} + \frac{(\alpha_2^{(n)})^2}{1-a_2^2}) - 2d\delta^{(n)}(\frac{\alpha_1^{(n)}}{a_1} log(1-a_1^2) + \frac{\alpha_2^{(n)}}{a_2} log(1-a_2^2)) + \\
(\delta^{(n)})^2 (\sum_{k=1}^{\infty} (\frac{a_1^k}{k})^2 + \sum_{k=1}^{\infty} (\frac{a_2^k}{k})^2)],
\end{gathered}$$

we can also rewrite the expression as follows, we put $\tau^{(n)} = (\alpha_1^{(n)} \quad \alpha_2^{(n)} \quad \delta^{(n)} \quad \delta^{(n)})$ and

$$\Gamma(\theta) = \begin{pmatrix}
\frac{d^2}{1-a_1^2} & 0 & -\frac{d}{a_1} log(1-a_1^2) & 0 \\
0 & \frac{d^2}{1-a_2^2} & 0 & -\frac{d}{a_2} log(1-a_2^2) \\
-\frac{d}{a_1} log(1-a_1^2) & 0 & \sum_{k=1}^{\infty} (\frac{a_1^k}{k})^2 & 0 \\
0 & -\frac{d}{a_2} log(1-a_2^2) & 0 & \sum_{k=1}^{\infty} (\frac{a_2^k}{k})^2
\end{pmatrix},$$

so, $(\lambda^{(n)})^2 = (\tau^{(n)})\Gamma_f(\theta)(\tau^{(n)})^\intercal$ with

$$\Gamma_f(\theta) = I(f)\sigma^2\Gamma(\theta),$$

then

$$2\sum_{M=0}^{N-1}(Z_{1,M,n}+Z_{2,M,n}) \longrightarrow \mathbb{N}(0,(\tau^{(n)})\Gamma_f(\theta)(\tau^{(n)})^{\intercal}).$$

Accordingly, for any period s, i.e., $i=1,...,s$, we have a $(2s\times 2s)$ symmetric matrix $\Gamma(\theta)$ of the form

$$\Gamma(\theta)=\begin{pmatrix}
\frac{d^2}{1-a_1^2} & 0 & \cdots & 0 & -\frac{d}{a_1}log(1-a_1^2) & 0 & \cdots & 0\\
0 & \frac{d^2}{1-a_2^2} & 0 & \vdots & 0 & -\frac{d}{a_2}log(1-a_2^2) & 0 & \vdots\\
\vdots & 0 & \ddots & 0 & \vdots & 0 & \vdots & 0\\
0 & \vdots & 0 & \frac{d^2}{1-a_s^2} & 0 & \vdots & 0 & -\frac{d}{a_s}log(1-a_s^2)\\
-\frac{d}{a_1}log(1-a_1^2) & 0 & \vdots & 0 & \sum_{k=1}^{\infty}(\frac{a_1^k}{k})^2 & 0 & \vdots & 0\\
0 & -\frac{d}{a_2}log(1-a_2^2) & 0 & \vdots & 0 & \sum_{k=1}^{\infty}(\frac{a_2^k}{k})^2 & 0 & \vdots\\
\vdots & 0 & \ddots & 0 & \vdots & 0 & \ddots & 0\\
0 & \cdots & 0 & \frac{d^2}{1-a_s^2} & 0 & \cdots & 0 & \sum_{k=1}^{\infty}(\frac{a_s^k}{k})^2,
\end{pmatrix},$$

and $\tau^{(n)}=(\alpha_1^{(n)},...,\alpha_s^{(n)},\delta^{(n)},...,\delta^{(n)})$, we get $(\lambda^{(n)})^2=(\tau^{(n)})\Gamma_f(\theta)(\tau^{(n)})^{\intercal}$ where $\Gamma_f(\theta)=I(f)\sigma^2\Gamma(\theta)$.

Under the regularity assumptions and for all sequences $(\tau^{(n)})$ such that $sup||\tau^{(n)}||<\infty$, when $N\longrightarrow\infty$, $2\sum_{M=0}^{N-1}\sum_{i=1}^{s}Z_{i,M,n}$ is **asymptotically normal** with **mean** 0 and **variance** equals $(\tau^{(n)})\Gamma_f(\theta)(\tau^{(n)})^{\intercal}$.

4.3.2 Local asymptotic minimaxity

In perspective of local asymptotic normality LAN$(\theta;\Gamma_f(\theta);\Delta_f^{(n)})$ property, this paragraph defines the **local Asymptotic Minimaxity** (LAM) criterion [24].
Let ℓ be a loss function from $\mathbb{R}^{s+1}$ to $\mathbb{R}_+$ such that $\ell(x)=\ell(-x),\forall x\in\mathbb{R}^{s+1}$, where the set $\{\ell\leq u\}$ is convex for each $u>0$.

Lemma 4.3.1 *Under $LAN(\theta; \Gamma_f(\theta); \Delta_f^{(n)})$, for all $\theta^{(n)}$, if*

$$\sqrt{n}(\theta^{(n)} - \theta) - \frac{[\Gamma(\theta)^{-1}]}{I(f)}\Delta_f^{(n)} = o_{p_\theta}(1), \tag{4.6}$$

then

i) $\theta^{(n)}$ is LAM under $H_f^{(n)}(\theta)$.

ii) If (4.6) is verified, $\theta^{(n)}$ is said to be $\theta-$regular.

Proof 16 *i) Under Kreiss's lemma 4.1 [66], the proof is inferred directly.*

ii) $\theta-$regularity concept.

4.3.3 Local asymptotic linearity

Since we have the local asymptotic quadratic (LAQ) form, we are able to derive the following lemma, which is considered to be a crucial concept in the construction of adaptive estimators for θ.

Lemma 4.3.2 *If $\theta^{(n)}$ is a sequence of estimators of θ such that $\sqrt{n}(\theta^{(n)} - \theta) = O_{p_\theta}(1)$, then*

$$\Delta_f^{(n)}(\theta^{(n)}) - \Delta_f^{(n)}(\theta) = -I(f)\Gamma(\theta)\tau^{(n)} + o_{p_\theta}(1), \tag{4.7}$$

Proof 17 *Since we have*

$$\Lambda_{f,\theta^{(n)}/\theta} = -\Lambda_{f,\theta/\theta^{(n)}} + o_{p_\theta}(1),$$

and

$$\Lambda_{f,\theta/\theta^{(n)}} = (\tau^{(n)})\Delta_f^{(n)}(\theta^{(n)}) - \frac{1}{2}(\tau^{(n)})\Gamma_f(\theta)(\tau^{(n)})^\intercal + o_{p_\theta}(1),$$

we conduct a quick calculation (see "Chapter 5 [94] for more details"), we obtain

$$\Delta_f^{(n)}(\theta) - \Delta_f^{(n)}(\theta + n^{-\frac{1}{2}}\tau^{(n)}) = \Gamma_f(\theta)(\tau^{(n)})^\intercal + o_{p_\theta}(1), \tag{4.8}$$

we consider $\theta^{(n)} = \theta + n^{-\frac{1}{2}}(n^{\frac{1}{2}}(\theta^{(n)} - \theta))$ and $\tau^{(n)} = n^{\frac{1}{2}}(\theta^{(n)} - \theta)$, we replace $\theta^{(n)}$ and $\tau^{(n)}$ by their expressions in (4.8), we get

$$\Delta_f^{(n)}(\theta) - \Delta_f^{(n)}(\theta^{(n)}) = \Gamma_f(\theta)n^{\frac{1}{2}}(\theta^{(n)} - \theta) + o_{p_\theta}(1).$$

4.4 Simulation

This section is devoted to a simulation study. We aim to verify the theoretical results of the local asymptotic normality.

We highlight the central sequence $\Delta_f^{(n)}$ that will be compared to the standard normal distribution. Note that $\Delta_f^{(n)} \in \mathbb{R}^{2s}$, so to reduce the complexity, we limit our study to the case of a period $s = 2$, i.e., $\Delta_f^{(n)} \in \mathbb{R}^4$.

We consider (4.4) and we deal with each component of $\Delta_f^{(n)}$ separately. For each component, we generate a $n = 2000$ central sequence with $a_1^{(n)} = 0.7 + n^{-\frac{1}{2}} \times 0.07$, $a_2^{(n)} = 0.92 + n^{-\frac{1}{2}} \times 0.09$ and $d^{(n)} = 0.49 + n^{-\frac{1}{2}} \times 0.04$ and we compare it to a n realization of a gaussian process with the same mean and variance of $\Delta_f^{(n)}$, where f is assumed to be normally distributed.

The QQ plot of each component is performed to check the local asymptotic normality of its distribution.

Furthermore, we consider other $a_i^{(n)}$ and $d^{(n)}$ values in the MSE tables provided with each figure to show the effect of varying them on the validity of the LAN property (see table i for component i; $i = 1, ..., 4$).

All figures are obtained for $n = 2000$.

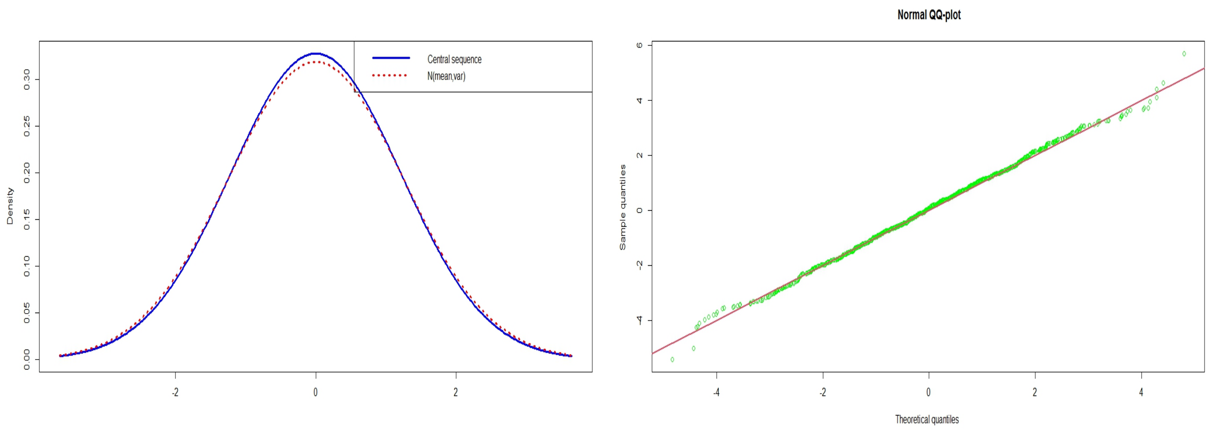

Figure 4.1: Density plot and QQ-plot of $\Delta_f^{(n)}$ first component

n	100	500	1000	2000
$a_1^{(n)} = 0.7 + n^{-\frac{1}{2}} \times 0.07$ $d^{(n)} = 0.49 + n^{-\frac{1}{2}} \times 0.04$	0.01491785	0.01338333	0.0003876789	4.228599e-05
$a_1^{(n)} = 0.58 + n^{-\frac{1}{2}} \times 0.04$ $d^{(n)} = 0.49 + n^{-\frac{1}{2}} \times 0.04$	0.03409093	0.0271874	0.01723482	0.007203508
$a_1^{(n)} = 0.3 + n^{-\frac{1}{2}} \times 0.025$ $d^{(n)} = 0.49 + n^{-\frac{1}{2}} \times 0.04$	0.7361413	0.6501581	0.2446102	0.2435163
$a_1^{(n)} = 0.7 + n^{-\frac{1}{2}} \times 0.07$ $d^{(n)} = 3.1 + n^{-\frac{1}{2}} \times 0.07$	1.191921	0.7517666	0.4540717	0.1869369
$a_1^{(n)} = 0.7 + n^{-\frac{1}{2}} \times 0.07$ $d^{(n)} = 0.08 + n^{-\frac{1}{2}} \times 0.001$	0.525892	0.2180738	0.09930603	0.03739237

Table 4.1: **MSE of $\Delta_f^{(n)}$ first component and the standard gaussian process**

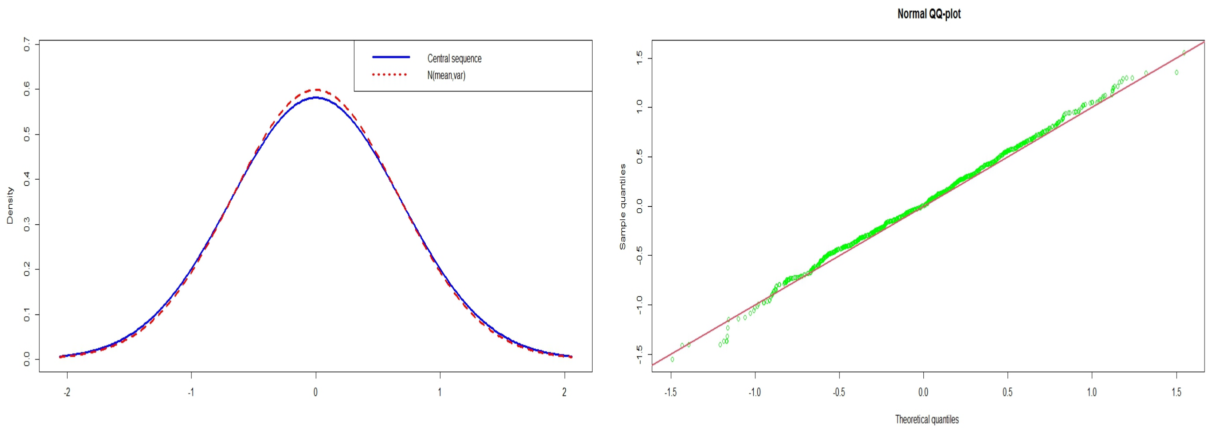

Figure 4.2: Density plot and QQ-plot of $\Delta_f^{(n)}$ second component

n	100	500	1000	2000
$a_2^{(n)} = 0.92 + n^{-\frac{1}{2}} \times 0.09$ $d^{(n)} = 0.49 + n^{-\frac{1}{2}} \times 0.04$	0.02352083	0.01206706	5.680078e-05	1.249709e-06
$a_2^{(n)} = 0.51 + n^{-\frac{1}{2}} \times 0.05$ $d^{(n)} = 0.49 + n^{-\frac{1}{2}} \times 0.04$	0.8573282	0.4648387	0.142002	0.1240862
$a_2^{(n)} = 0.13 + n^{-\frac{1}{2}} \times 0.01$ $d^{(n)} = 0.49 + n^{-\frac{1}{2}} \times 0.04$	7.10016	5.446648	1.674653	1.555713
$a_2^{(n)} = 0.92 + n^{-\frac{1}{2}} \times 0.09$ $d^{(n)} = 3.1 + n^{-\frac{1}{2}} \times 0.07$	0.9153858	0.6422962	0.3355191	0.08703649
$a_2^{(n)} = 0.92 + n^{-\frac{1}{2}} \times 0.09$ $d^{(n)} = 0.08 + n^{-\frac{1}{2}} \times 0.001$	0.1367081	0.1352859	0.09448176	0.01428824

Table 4.2: **MSE of $\Delta_f^{(n)}$ second component and the standard gaussian process**

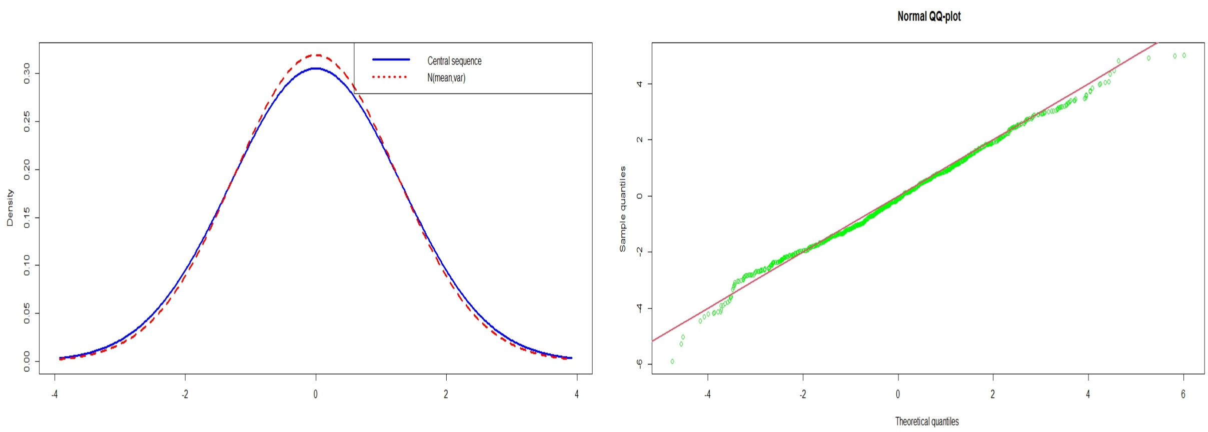

Figure 4.3: Density plot and QQ-plot of $\Delta_f^{(n)}$ third component

n	100	500	1000	2000
$a_1^{(n)} = 0.7 + n^{-\frac{1}{2}} \times 0.07$ $d^{(n)} = 0.49 + n^{-\frac{1}{2}} \times 0.04$	0.02324896	0.01257646	0.0009121133	4.138632e-05
$a_1^{(n)} = 0.58 + n^{-\frac{1}{2}} \times 0.04$ $d^{(n)} = 0.49 + n^{-\frac{1}{2}} \times 0.04$	0.284103	0.1585792	0.02508525	0.005351969
$a_1^{(n)} = 0.3 + n^{-\frac{1}{2}} \times 0.025$ $d^{(n)} = 0.49 + n^{-\frac{1}{2}} \times 0.04$	1.6865	1.140523	0.8553154	0.490919
$a_1^{(n)} = 0.7 + n^{-\frac{1}{2}} \times 0.07$ $d^{(n)} = 3.1 + n^{-\frac{1}{2}} \times 0.07$	1.280266	0.9940527	0.78614	0.1801243
$a_1^{(n)} = 0.7 + n^{-\frac{1}{2}} \times 0.07$ $d^{(n)} = 0.08 + n^{-\frac{1}{2}} \times 0.001$	0.7541451	0.4036433	0.09617421	0.06160828

Table 4.3: **MSE of $\Delta_f^{(n)}$ third component and the standard gaussian process**

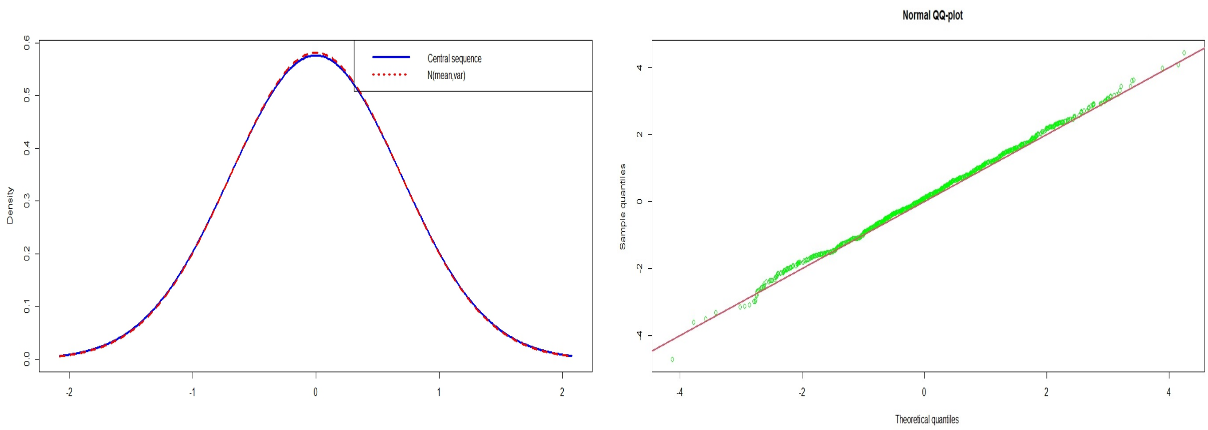

Figure 4.4: Density plot and QQ-plot of $\Delta_f^{(n)}$ fourth component

n	100	500	1000	2000
$a_2^{(n)} = 0.92 + n^{-\frac{1}{2}} \times 0.09$ $d^{(n)} = 0.49 + n^{-\frac{1}{2}} \times 0.04$	0.1024018	0.03029323	0.0005531046	5.813622e-06
$a_2^{(n)} = 0.51 + n^{-\frac{1}{2}} \times 0.05$ $d^{(n)} = 0.49 + n^{-\frac{1}{2}} \times 0.04$	1.062268	1.207582	0.8939105	0.526109
$a_2^{(n)} = 0.13 + n^{-\frac{1}{2}} \times 0.01$ $d^{(n)} = 0.49 + n^{-\frac{1}{2}} \times 0.04$	9.776774	6.974689	2.684501	1.75958
$a_2^{(n)} = 0.92 + n^{-\frac{1}{2}} \times 0.09$ $d^{(n)} = 3.1 + n^{-\frac{1}{2}} \times 0.07$	1.299422	1.070233	0.421268	0.1229311
$a_2^{(n)} = 0.92 + n^{-\frac{1}{2}} \times 0.09$ $d^{(n)} = 0.08 + n^{-\frac{1}{2}} \times 0.001$	0.1384956	0.1136398	0.05685105	0.04381796

Table 4.4: **MSE of $\Delta_f^{(n)}$ fourth component and the standard gaussian process**

In the figures 4.1, 4.2, 4.3 and 4.4, the red dotted lines represent the density curves of the normal distribution and the blue lines are the density curves of the central sequence components. The closeness between all pairs of curves allows us to deduce that the four components have a normal distribution. Thus, $\Delta_f^{(n)}$ has a normal distribution.
The figures also show the normal QQ plots of $\Delta_f^{(n)}$ components, which are almost aligned with the straight diagonal lines of the gaussian processes, indicating that $\Delta_f^{(n)}$ has a normal limit distribution.

For different sample sizes and different parameter values, we calculate the mean squared error (MSE) between the gaussian processes and the generated $\Delta_f^{(n)}$ components. The findings were presented in tables 4.1, 4.2, 4.3 and 4.4.

- The MSE decreases as the sample size n increases.
- The MSE is negligible when $a_i^{(n)}$ $(i = 1, 2)$ approaches 1.

- The MSE increases significantly as $a_i^{(n)}$ $(i = 1, 2)$ decreases and approaches zero.
- The MSE decreases as $d^{(n)}$ approaches to $\frac{1}{2}$, becomes small as $d^{(n)}$ approaches 0 and increases as $d^{(n)}$ increases significantly.

A low MSE value means that the distribution of the central sequence is close to the standard normal distribution with the same mean and variance.

4.5 Conclusion

In this chapter, we have established the LAN property of the periodic fractional autoregressive model, where we have shown the effect of periodicity on the LAN property, especially on the form of the central sequence, which corresponds to the period s, and then used it to investigate the existence of LAM estimation through the local asymptotic linearity property.
The theoretical LAN results were illustrated through a simulation study.

GENERAL CONCLUSION AND PERSPECTIVES

All in all, in this study, motivated by a meteorological dataset, we developed a time series model that integrates both long- and short-memory behavior with periodic patterns. We began by presenting the main probabilistic properties, such as the conditions for causality and invertibility, and demonstrated through the covariance function that the proposed model is non-stationary; rather, it belongs to the class of periodically correlated models.

We then derived the autocovariance and autocorrelation functions together with their asymptotic behavior and proposed an estimation method based on minimizing the distance between two probability densities. The asymptotic properties of the resulting estimator were also established. Furthermore, we compared this estimation approach with an analog of the maximum likelihood estimator, namely the conditional sum of squares (CSS) estimator. Finally, we proved that the model satisfies the local asymptotic normality (LAN) property, which allowed us to establish its local asymptotic linearity and local asymptotic minimaxity (LAM).

All the theoretical results were thoroughly examined through extensive simulation studies, where several scenarios and parameter configurations were considered. The obtained results confirmed the validity and robustness of the proposed theoretical findings.

Overall, this work provides a comprehensive probabilistic and statistical framework for modeling periodic time series with both long- and short-memory components. However, several aspects remain open for further investigation. In particular, the theoretical construction of confidence and prediction intervals could be explored to provide a more complete inferential framework. Likewise, a deeper asymptotic study of the CSS estimator would shed further light on its efficiency and practical relevance.
Another natural extension would be the development of periodic fractional autoregressive models with mixed or dependent errors, as well as the design of adaptive tests suitable for such settings. Finally, an important future step consists in applying the proposed methodology to real data, possibly through hybrid approaches involving fractional autoregressive models.

These directions form the main perspectives of this work, which we believe offer promising opportunities for both theoretical and applied developments in periodic time series analysis.

BIBLIOGRAPHY

[1] Abramowitz, M., and Stegun, I. A. *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, vol. 55. US Government Printing Office, 1968.

[2] Ahmed, I. Detection of nonlinearity and stochastic nature in time series by delay vector variance method. *Int. J. Eng. Technol., Int. J. Eng. Sci 10*, 2 (2010), 22–27.

[3] Amimour, A., and Belaide, K. Local asymptotic normality for a periodically time varying long memory parameter. *Communications in Statistics-Theory and Methods 51*, 9 (2022), 2936–2952.

[4] Amimour, A., Belaide, K., and Hili, O. Minimum hellinger distance estimates for a periodically time-varying long memory parameter. *Comptes Rendus. Mathématique 360*, G10 (2022), 1153–1162.

[5] Atkins, F. J., and Chen, J. Some statistical properties of deregulated electricity prices in alberta. *University of Calgary Department of Economics Discussion Paper 2002 6* (2002).

[6] Baillie, R. T. Long memory processes and fractional integration in econometrics. *Journal of econometrics 73*, 1 (1996), 5–59.

[7] Bartlett, M. S. On the theoretical specification and sampling properties of autocorrelated time-series. *Supplement to the Journal of the Royal Statistical Society 8*, 1 (1946), 27–41.

[8] Beran, R. Minimum hellinger distance estimates for parametric models. *The annals of Statistics* (1977), 445–463.

[9] Boshnakov, G. Periodically correlated sequences: some properties and recursions. Lule University of Technology (1994).

[10] Box, G. E., Jenkins, G. M., Reinsel, G. C., and Ljung, G. M. *Time series analysis: forecasting and control.* John Wiley & Sons, 2015.

[11] Brockwell, P. J. *Time series: Theory and methods.* Springer-Verlag, 1991.

[12] Brookes, B. Stochastic processes. by jl doob. pp. vii, 654. 80s. 1953.(new york: Wiley. london: Chapman and hall). *The Mathematical Gazette 38*, 325 (1954), 236–238.

[13] Brown, R. G. *Smoothing, forecasting and prediction of discrete time series.* Courier Corporation, 2004.

[14] Cao, L. Support vector machines experts for time series forecasting. *Neurocomputing 51* (2003), 321–339.

[15] Cheung, Y.-W. Long memory in foreign-exchange rates. *Journal of Business & Economic Statistics 11*, 1 (1993), 93–101.

[16] Cheung, Y.-W., and Lai, K. S. A search for long memory in international stock market returns. *Journal of International Money and Finance 14*, 4 (1995), 597–615.

[17] Choy, K., Hallin, M., Serroukh, A., and Taniguchi, M. Local asymptotic normality for regression models with long-memory disturbance. *The Annals of Statistics 27*, 6 (1999), 2054–2080.

[18] Chung, C.-F. Estimating a generalized long memory process. *Journal of econometrics 73*, 1 (1996), 237–259.

[19] Cipra, T. Periodic moving average process. *Aplikace matematiky 30*, 3 (1985), 218–229.

[20] Diongue, A. K. *Modélisation longue mémoire multivariée: applications aux problématiques du producteur d'EDF dans le cadre de la libéralisation du marché européen de l'électricité*. PhD thesis, École normale supérieure de Cachan-ENS Cachan, 2005.

[21] Durrett, R. *Probability: theory and examples*, vol. 49. Cambridge university press, 2019.

[22] Einstein, A., et al. On the motion of small particles suspended in liquids at rest required by the molecular-kinetic theory of heat. *Annalen der physik 17*, 549-560 (1905), 208.

[23] Engle, R. F., and Granger, C. W. Co-integration and error correction: representation, estimation, and testing. *Econometrica: journal of the Econometric Society* (1987), 251–276.

[24] Fabian, V., and Hannan, J. On estimation and adaptive estimation for locally asymptotically normal families. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete 59*, 4 (1982), 459–478.

[25] Ferrara, L., and Guegan, D. Forecasting financial times series with generalized long memory processes. *Advances in quantitative asset management* (2000), 319–342.

[26] Garel, B., and Hallin, M. Local asymptotic normality of multivariate arma processes with a linear trend. *Annals of the Institute of Statistical Mathematics 47* (1995), 551–579.

[27] Garel, B., and Hallin, M. Local asymptotic normality of multivariate arma processes with a linear trend. *Annals of the Institute of Statistical Mathematics 47* (1995), 551–579.

[28] Gautama, T., Mandic, D. P., and Van Hulle, M. M. The delay vector variance method for detecting determinism and nonlinearity in time series. *Physica D: Nonlinear Phenomena 190*, 3-4 (2004), 167–176.

[29] Gelfand, A. E., and Smith, A. F. Sampling-based approaches to calculating marginal densities. *Journal of the American statistical association 85*, 410 (1990), 398–409.

[30] Ghosh, K. Time series analysis: A brief history and its future challenges. *Indian Science Cruiser 34*, 5 (2020), 22–27.

[31] Gladyshev, E. Periodically correlated random sequences. In *Doklady Akademii Nauk* (1961), vol. 137, Russian Academy of Sciences, pp. 1026–1029.

[32] Gladyshev, E. Periodically and almost-periodically correlated random processes with a continuous time parameter. *Theory of Probability & Its Applications 8*, 2 (1963), 173–177.

[33] Gonçalves, E. Une généralisation des processus arma. *Annales d'Economie et de Statistique 5* (1987), 109–145.

[34] Gonçalves, E. *Processus fractionnaires.* PhD thesis, Lille 1, 1988.

[35] Granger, C. W. Investigating causal relations by econometric models and cross-spectral methods. *Econometrica: journal of the Econometric Society* (1969), 424–438.

[36] Granger, C. W., and Andersen, A. On the invertibility of time series models. *Stochastic processes and their applications 8*, 1 (1978), 87–92.

[37] Granger, C. W., and Joyeux, R. An introduction to long-memory time series models and fractional differencing. *Journal of time series analysis 1*, 1 (1980), 15–29.

[38] Grattan-Guinness, I. *Landmark writings in Western mathematics 1640-1940*. Elsevier, 2005.

[39] Grenander, U., and Rosenblatt, M. *Statistical analysis of stationary time series*, vol. 320. American Mathematical Soc., 2008.

[40] Guegan, D. Processus à longue mémoire. propriétés probabilistes et statistiques. In *Annales de l'ISUP* (1991), vol. 36, pp. 5–41.

[41] Haddad, S., and Belaide, K. Local asymptotic normality for long-memory process with strong mixing noises. *Communications in Statistics-Theory and Methods 49*, 12 (2020), 2817–2830.

[42] Hájek, J. Local asymptotic minimax and admissibility in estimation. In *Proceedings of the sixth Berkeley symposium on mathematical statistics and probability* (1972), vol. 1, pp. 175–194.

[43] Hájek, J., and Sidak, J. Theory of rank tests new york: Academic.

[44] Hannan, E. *Time Series*. Analysis, Methuen, London, 1960.

[45] Harvey, A. C. Forecasting, structural time series models and the kalman filter (1990).

[46] Higuchi, T. Approach to an irregular time series on the basis of the fractal theory. *Physica D: Nonlinear Phenomena 31*, 2 (1988), 277–283.

[47] Higuchi, T. Relationship between the fractal dimension and the power law index for a time series: a numerical investigation. *Physica D: Nonlinear Phenomena 46*, 2 (1990), 254–264.

[48] Hili, O. On the estimation of nonlinear time series models. *Stochastics: An International Journal of Probability and Stochastic Processes 52*, 3-4 (1995), 207–226.

[49] Holt, C. C. Forecasting trends and seasonals by exponentially weighted moving averages. *ONR Memorandum 52*, 52 (1957), 5–10.

[50] Hosking, J. Fractional differencing. *Biometrika 68*, 1 (1981), 165–176.

[51] Hosking, J. R. Modeling persistence in hydrological time series using fractional differencing. *Water resources research 20*, 12 (1984), 1898–1908.

[52] Hurd, H. L. *An investigation of periodically correlated stochastic processes.* Duke University, 1970.

[53] Hurst, H. E. Long-term storage capacity of reservoirs. *Transactions of the American society of civil engineers 116*, 1 (1951), 770–799.

[54] Iacus, S. M., et al. *Simulation and inference for stochastic differential equations: with R examples*, vol. 486. Springer, 2008.

[55] Jo, T. C. The effect of virtual term generation on the neural-based approaches to time series prediction. In *2003 4th International Conference on Control and Automation Proceedings* (2003), pp. 516–520.

[56] Jochmann, H. Period variations of the chandler wobble. *Journal of Geodesy 77* (2003), 454–458.

[57] Jones, R. H. Maximum likelihood fitting of arma models to time series with missing observations. *Technometrics 22*, 3 (1980), 389–395.

[58] Kakutani, S. On equivalence of infinite product measures. *Annals of Mathematics 49*, 1 (1948), 214–224.

[59] Kalman, R. E. A new approach to linear filtering and prediction problems (1960).

[60] Kalman, R. E., and Bucy, R. S. New results in linear filtering and prediction theory (1961), 95-108.

[61] Karr, A. F. *Probability.* Springer-Verlag, New York, 1993.

[62] Kendall, M. G. On the analysis of oscillatory time-series. *Journal of the Royal Statistical Society 108*, 1/2 (1945), 93–141.

[63] KITAGAWA, G., AND GERSCH, W. *Smoothness priors analysis of time series*, vol. 116. Springer Science & Business Media, 2012.

[64] KOLMOGOROV, A. N., AND BHARUCHA-REID, A. T. *Foundations of the theory of probability: Second English Edition.* Courier Dover Publications, 2018.

[65] KRAFT, C. Some conditions for consistency and uniform consistency of statistical procedures. *University of California Publication in Statistics 2* (1955), 125–141.

[66] KREISS, J.-P. On adaptive estimation in autoregressive models when there are nuisance functions. *Statistics & Risk Modeling 5*, 1-2 (1987), 59–76.

[67] KREISS, J.-P. On adaptive estimation in stationary arma processes. *The Annals of Statistics* (1987), 112–133.

[68] LAURITZEN, S. L. Time series analysis in 1880: A discussion of contributions made by tn thiele. *International Statistical Review/Revue Internationale de Statistique* (1981), 319–331.

[69] LAURITZEN, S. L. Aspects of tn thiele's contributions to statistics. *Bulletin of the International Statistical Institute 58* (1999), 27–30.

[70] LE CAM, L. Locally asymptotically normal families of distributions. certain approximations to families of distributions and their use in the theory of estimation and testing hypotheses. *Univ. California Publ. Statist. 3* (1960), 37.

[71] LE CAM, L. *Asymptotic methods in statistical decision theory.* Springer Science & Business Media, 2012.

[72] LE CAM, L. M., AND YANG, G. L. *Asymptotics in statistics: some basic concepts.* Springer Science & Business Media, 2000.

[73] LI, W. K., AND MCLEOD, A. I. Fractional time series modelling. *Biometrika 73*, 1 (1986), 217–221.

[74] Matusita, K. Decision rules, based on the distance, for problems of fit, two samples, and estimation. *The Annals of Mathematical Statistics* (1955), 631–640.

[75] Mbeke, K. S., and Hili, O. Estimation of a stationary multivariate arfima process. *Afrika Statistika 13*, 3 (2018), 1717–1732.

[76] Mbeke, Kévin Stanislas, y. s. *Statistique Paramétrique des Processus Stationnaires Longue Mémoire*. PhD thesis.

[77] Muth, J. F. Optimal properties of exponentially weighted forecasts. *Journal of the american statistical association 55*, 290 (1960), 299–306.

[78] Ndongo, M., Diongue, A. K., Diop, A., and Hili, O. Estimation of fractional arima process with stable innovations: A monte carlo study (2015).

[79] Nielsen, A. *Practical time series analysis: Prediction with statistics and machine learning*. O'Reilly Media, 2019.

[80] Odaki, M. On the invertibility of fractionally differenced arima processes. *Biometrika 80*, 3 (1993), 703–709.

[81] Ogata, Y., and Abe, K. Some statistical features of the long-term variation of the global and regional seismic activity. *International Statistical Review/Revue Internationale de Statistique* (1991), 139–161.

[82] Palma, W. *Long-memory time series: theory and methods*. John Wiley & Sons, 2007.

[83] Palma, W. *Time series analysis*. John Wiley & Sons, 2016.

[84] Quenouille, M. H. Associated measurements (1957).

[85] Quenouille, M. H. The analysis of multiple time-series. *(London:Griffin)* (1957).

[86] Rao, B. P. *Nonparametric functional estimation*. Academic press, 2014.

[87] Rao, J., and Tintner, G. On the variate difference method. *Australian Journal of Statistics 5*, 3 (1963), 106–116.

[88] Reisen, V. A., and Lopes, S. Some simulations and applications of forecasting long-memory time-series models. *Journal of Statistical Planning and Inference 80*, 1-2 (1999), 269–287.

[89] Ross, S. M. *Stochastic processes.* John Wiley & Sons, 1995.

[90] Roussas, G. G. Asymptotic distribution of the log-likelihood function for stochastic processes. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete 47*, 1 (1979), 31–46.

[91] Roy, M., Kumar, V. R., Kulkarni, B., Sanderson, J., Rhodes, M., and Stappen, M. v. Simple denoising algorithm using wavelet transform. *arXiv preprint nlin/0002028* (2000).

[92] Seba, D., and Belaide, K. On several local asymptotic properties for fractional autoregressive models with strong mixing noises. *Communications in Statistics-Simulation and Computation 53*, 4 (2024), 1744–1757.

[93] Seba, D., Benaklef, N., Belaide, K., et al. Time series analysis for modeling and predicting confirmed cases of influenza a in algeria. *Russian Journal of Infection and Immunity 15*, 1 (2025), 168–172.

[94] Serroukh, A. *Inférence asymptotique paramétrique et non paramétrique pour les modèles ARMA fractionnaires.* PhD thesis, Institut Statist., Univ. Libre Bruxelles, 1996.

[95] Slutzky, E. The summation of random causes as the source of cyclic processes. *Econometrica: Journal of the Econometric Society* (1937), 105–146.

[96] Smoluchowski, M. The kinetic theory of brownian molecular motion and suspensions. *Ann. Phys 21* (1906), 756–780.

[97] Soares, L. J., and Souza, L. R. Forecasting electricity demand using generalized long memory. *International Journal of Forecasting 22*, 1 (2006), 17–28.

[98] Swensen, A. R. The asymptotic distribution of the likelihood ratio for autoregressive time series with a regression trend. *Journal of Multivariate Analysis 16*, 1 (1985), 54–70.

[99] Tamura, R. N., and Boos, D. D. Minimum hellinger distance estimation for multivariate location and covariance. *Journal of the American Statistical Association 81*, 393 (1986), 223–229.

[100] Tintner, G. *The variate difference method.* Principia Press Bloomington, IN, 1940.

[101] Troutman, B. M. Some results in periodic autoregression. *Biometrika 66*, 2 (1979), 219–228.

[102] Tsay, R. S. Time series and forecasting: Brief history and future research. *Journal of the American Statistical Association 95*, 450 (2000), 638–643.

[103] Van der Vaart, A. W. *Asymptotic statistics*, vol. 3. Cambridge university press, 2000.

[104] WA, G. Cyclostationarity in communications and signal processing. *IEEE Press.* (1994).

[105] Walker, G. T. On periodicity in series of related terms. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character 131*, 818 (1931), 518–532.

[106] Wang, W.-C., Chau, K.-W., Cheng, C.-T., and Qiu, L. A comparison of performance of several artificial intelligence methods for forecasting monthly discharge time series. *Journal of hydrology 374*, 3-4 (2009), 294–306.

[107] Winters, P. R. Forecasting sales by exponentially weighted moving averages. *Management science 6*, 3 (1960), 324–342.

[108] Wold, H. *A study in the analysis of stationary time series.* PhD thesis, Almqvist & Wiksell, 1938.

[109] Wu, W. B., and Mielniczuk, J. Kernel density estimation for linear processes. *The Annals of statistics 30*, 5 (2002), 1441–1459.

[110] Xiang, S. Minimum hellinger distance estimation in a semiparametric mixture model (2012).

[111] Yaglom, A. The correlation theory of processes whose nth difference constitute a stationary process. *Matem. sb 37*, 79 (1955), 141.

[112] Yule, G. On a method of investigating periodicities in disturbed series with a special reference to wolfer s sunspot numbers, philosphical transactions, 1927, 226a. *Rs. crore IPFS Rs. crore IPFS* (1927).

[113] Yule, G. U. Why do we sometimes get nonsense-correlations between time-series?–a study in sampling and the nature of time-series. *Journal of the royal statistical society 89*, 1 (1926), 1–63.

[114] Zadeh, L. A., and Ragazzini, J. R. An extension of wiener's theory of prediction. *Journal of Applied Physics 21*, 7 (1950), 645–655.

Résumé

Dans cette thèse, nous avons étudié les modèles autorégressifs fractionnaires d'ordre 1 à coefficients périodiques. Ces modèles généralisent les modèles autorégressifs fractionnaires d'ordre 1 classiques à coefficients constants, qui ne permettent pas de capturer le caractère de périodicité présent dans divers phénomènes pratiques. Dans un premier temps, nous avons examiné les principales propriétés probabilistes, telles que les conditions de causalité et d'inversibilité, la stationnarité, ainsi que le comportement asymptotique des fonctions d'autocovariance et d'autocorrélation, qui caractérisent entièrement un processus stochastique. Dans un second temps, nous nous sommes intéressés à l'estimation des paramètres. Comme le modèle étudié est non stationnaire mais périodiquement corrélé, les méthodes classiques d'estimation ne sont pas adaptées. Pour cette raison, nous avons proposé deux techniques d'estimation des paramètres du modèle considéré et étudié la consistance des estimateurs obtenus. La dernière partie de la thèse est consacrée à l'étude des propriétés asymptotiques locales. Nous avons établi la propriété de normalité asymptotique locale, la linéarité asymptotique locale et l'existence d'un estimateur minimax. Enfin, tous les résultats principaux présentés dans cette thèse ont été largement vérifiés et validés à travers des études de simulation.

Mots clés: Processus autorégressif fractionnaire, périodicité, auto-covariance, estimation, distance de Hellinger, estimateur des moindres carrés conditionnels, normalité asymptotique locale, linéarité asymptotique locale.

Abstract

In this thesis, we studied first-order fractional autoregressive models with periodic coefficients. These models generalize the classical first-order fractional autoregressive models with constant coefficients, which fail to capture the periodicity observed in many practical phenomena. First, we examined the main probabilistic properties, such as the conditions for causality and invertibiliy, stationarity, and the asymptotic behavior of the autocovariance and autocorrelation functions, which fully characterize a stochastic process. Next, we focused on parameter estimation. Since the studied model is non-stationary but periodically correlated, classical estimation methods are not suitable. Therefore, we proposed two techniques for estimating the parameters of the considered model and investigated the consistency of the resulting estimators. The final part of the thesis is devoted to the study of local asymptotic properties. We established the local asymptotic normality property, local asymptotic linearity, and the existence of a minimax estimator. Finally, all the main theoretical results presented in this thesis were thoroughly examined and validated through extensive simulation studies.

Key words: Fractional autoregressive process, periodicity, autocovariance, estimation, Hellinger distance, conditional least squares estimator, local asymptotic normality, local asymptotic linearity, simulation.

ملخص

في هذه الأطروحة، قمنا بدراسة نماذج الانحدار الذاتي الكسري من الدرجة الأولى ذات المعاملات الدورية. تعمم هذه النماذج نماذج الانحدار الذاتي الكسري الكلاسيكية من الدرجة الأولى ذات المعاملات الثابتة، والتي لا تستطيع التقاط الطبيعة الدورية الموجودة في العديد من الظواهر العملية. في البداية، قمنا بدراسة الخصائص الاحتمالية الرئيسية، مثل شروط السببية والقابلية للعكس، والثبات، والسلوك المقارب لدوال التغاير الذاتي والارتباط الذاتي، والتي تميز بشكل كامل عملية عشوائية. بعد ذلك، ركزنا على التقدير. وبما أن النموذج المدروس غير ثابت ولكنه مترابط بشكل دوري، فإن طرق التقدير الكلاسيكية ليست مناسبة. لذلك، قمنا باقتراح تقنيتين لتقدير معاملات النموذج المدروس ودراسة اتساق المقدرات التي تم الحصول عليها. الجزء الأخير من الأطروحة مخصص لدراسة الخصائص المقاربية المحلية. قمنا بإثبات خاصية الانتظام المقارب المحلي، والخطية المقاربية المحلية، ووجود مقدر ذو حد ادنى-اعلى.

تم التحقق من جميع النتائج الرئيسية المقدمة في هذه الأطروحة بشكل مكثف والتحقق من صحتها عبر دراسات المحاكاة.

الكلمات المفتاحية : عملية الانحدار الذاتي الكسري, دورية, التغاير الذاتي, تقدير, مسافة هيلنجر, مقدر المربعات الصغرى الشرطي, التقارب الطبيعي المحلي, التقارب الخطي المحلي, محاكاة.